Psychometrika

VOLUME VIII—1943
JANUARY-DECEMBER

PUBLISHED QUARTERLY
By THE PSYCHOMETRIC SOCIETY
AT 23 WEST COLORADO AVENUE
COLORADO SPRINGS, COLORADO

Psychometrika

CONTENTS

VOLUME EIGHT MARCH 1943 NUMBER ONE

PSYCHOMETRIKA—VOL. 8, NO. 1
MARCH, 1943

A GENERAL THEORY OF LEARNING AND CONDITIONING: PART I

WALTER PITTS
THE UNIVERSITY OF CHICAGO

In the first of two parts in which a general mathematical theory of non-symbolic learning and conditioning is constructed, the sections of the theory dealing with non-symbolic learning and conditioning are presented, and a number of its qualitative implications are compared with available experimental results. In general, the agreement is found to be rather close.

1. *Introduction: the current theories.*

The field of conditioning and learning has attained a development on the purely experimental side which renders it an excellent point for the entry of quantitative theory into psychology; and the exploitation of various aspects of this domain for theoretical purposes has already made a promising beginning. The work of this kind done so far has been chiefly from two standpoints: the first, exemplified in the researches of Householder, Landahl, and Rashevsky, as well as others of the same group (summarized in **39** and **40**), attempts with some success to explain the phenomena directly upon a neurological basis, while the second, comprising studies such as those of Gullikson (**15**), Gullikson and Wolfle (**16**), Hull (**27**, **28**, **29**), and Spence (**48**), prefers to elaborate first a macroscopic account of behavior per se, while leaving the neurological foundations until a later stage. These two approaches are of course rather complementary than competitive: the development of theoretical neurology provides very many suggestions for macroscopic work, and the latter simplifies the neurological problem by requiring mechanisms to account for only a few general propositions instead of a multitude of facts in no obvious relation.

So far, the theoretical structures built according to this second viewpoint have shared the common defect of being in a certain sense incompletely quantified. To a considerable degree the stimuli and responses involved, and the successive trials in the learning process, are regarded as unanalysed units; and the laws of variation for the strengths of tendencies to make given responses are stated in terms

of the *number* of previous learning trials belonging to specified categories. This procedure, to be sure, has advantages in intuitive and mathematical simplicity which are not to be despised; but it nevertheless conduces to a certain lack of generality. In particular, if we are to count the trials in the learning process, classifying them separately only into a few groups—such as temporal order, or status as success or failure—we must arrange them in a predetermined temporal order and make them as much similar *inter se* as is experimentally feasible, so that in even so slight a modification of the experimental routine as spacing the learning trials irregularly, or varying some of the characteristics of the learning situation among the trials, we can no longer make a theoretical prediction at all. Moreover, there is usually no obvious way of generalizing a system constructed on these lines to such slightly different cases, while for more radically altered situations the difficulty is correspondingly worse. To remedy this, we shall seek to analyze the learning process in a more detailed fashion than before, so that we may state the contribution of each trial to learning in a way depending upon its relevant characteristics and those of the previous trials; and then to essay a statement of our ideas in a form more completely accessible to the methods of mathematical analysis, which, for their most fruitful use, enjoin substantial continuity in the quantities considered. In this way we may use the same experimental generalizations upon which the preceding theories have been based to construct a new one, similar to them in many respects, but rather more extensive in the scope and detail of its considerations. More precisely, we shall consider the results of our discussion applicable to all aspects of learning and conditioning in which the effect of symbolic or verbal factors is not of great significance; and within this field we shall deal with all cases of learning and conditioning in which independent or related stimuli, with given original tendencies to produce specified types of response, are distributed over time in specified intensities in an arbitrary way, continuous or otherwise; and in which affective stimulation, if this form part of the experimental routine, is distributed in any given manner.

In partial confirmation of our hypotheses, we shall point out how most of the principal experimental generalizations can be inferred from the theory, at least as regards comparative order of magnitude; while a rigid quantitative test would require data in a detail not ordinarily given in experimental results. The theory does not seem too difficult to verify in most of its aspects, however, by a fairly extended and precise set of experiments, whose performance would also provide direct information upon a number of matters of considerable

import upon which little data are available, and which, even if disconfirming our present system in some of its aspects, would assuredly make suggestions leading to a better one of comparable range and generality.

The presentation of the theory will be divided into two parts. Part I deals only with the case where the stimuli and responses are wholly independent, so that transfer and generalization do not occur, and proposes a law of variation for the reaction-tendency which takes into account all of classical conditioning and the various sorts of inhibition affecting it. Part II extends the discussion, still under the hypothesis of complete independence, to cases where reward and punishment are involved as motivating factors, then generalizes all the preceding results to the case where the stimuli and the responses are related psychophysically, thus providing a theory of transfer, generalization, and discrimination, and concludes with a more precise statement of the observational interpretation of the terms and quantities we shall employ.

2. *The determination of the simple response-tendency.*

As mentioned above, we shall suppose in this part that the stimuli are completely unlike, so that generalization and discrimination do not occur; that the reactions are very much different, so that transfer does not arise; and that the reaction-thresholds do not influence one another, so that there is no "generalization of inhibition." Although in most cases this is not entirely realized, in many types of experiment which will readily occur to the reader, such as classical conditioning for one, and serial learning of properly selected nonsense-syllables for another, it is perhaps sufficiently so for practice. In any case, a prior development of the theory for the independent case, before introducing the various complications involved in related quantities, has an heuristic and intuitive value which will presently become apparent.

We shall begin by introducing the notion of the *simple response tendency.* In the absence of inhibition, a reaction of the given sort will occur whenever this quantity exceeds its reaction-threshold; in more complicated circumstances the occurrence of the reaction depends upon the supraliminal character of some quantity which, regarded as a function of the simple response tendency alone, is linear, though in fact depending on other variables as well. A more complete definition of this term will be given in the last section: for the present, the foregoing may supply the reader with sufficient intuitive understanding to follow the subsequent arguments. If the different re-

action-tendencies involved in a situation be assigned numbers from 1 to M, we shall denote the simple response-tendency, for the j-th reaction, evaluated at the time t, by $E_j(t)$. If there be N different types of stimulus involved in the situation, we may assign numbers to them also, say from 1 to N. We shall then denote by $P_i(t)$ $[i = 1, 2, \cdots, N]$ the logarithm of intensity of the type i stimulation being presented at the time t.

In general, the presentation of type i stimulation at a given time will generate an unconditioned simple response-tendency of type j over a certain period of time thereafter, for suitable i and j. If a unit intensity of type i stimulation be applied at a time t, then the resulting *unconditioned* response-tendency of type j, effective at a later time $t + \delta$, will be denoted by $T^0_{ij}(\delta, t)$. We may define this quantity more exactly as follows: let the average value of P_i throughout a small interval Δt about t be $P_i(t)$; let the average value of E_j during a small interval $\Delta' t$ about $t + \delta$ be $\bar{E}_j(t + \delta)$; let P_i vanish outside Δt, and P_k, $k \neq i$, vanish everywhere; and suppose that no conditioning of the stimulus-type i to the reaction j has occurred. Then the limit $\bar{E}_j(t + \delta)/P_i(t)$ as Δt and $\Delta' t$ approach zero about t and $t + \delta$, respectively, is $T^0_{ij}(\delta, t)$. If in this definition we omit the requirement that type i stimulation has not been conditioned to the reaction j, so that we consider the gross magnitude of the response-tendency E_j, we shall obtain a function to be denoted by $T_{ij}(\delta, t)$.

Supposing that stimuli of type i have been applied over the period of time $t = 0$ to $t = t$, we proceed to calculate the total unconditioned tendency to make a type j response at time t which results from this. If the various increments in E_j emanating from different previous intervals combine additively, and the relation between the strength of the stimulus and the magnitude of the elicited reaction-tendency is the familiar logarithmic one, as found, for example, by Kupalov and Gantt (34), this tendency will be, by definition,

$$\int_0^t P_i(\vartheta) T^0_{ij}(t - \vartheta, \vartheta)\, d\vartheta .$$

Combining such quantities for all stimulus-types, and remembering our independence assumption, we shall obtain the total unconditioned part of $E_j(t)$, which is

$$S_j(t) = \int_0^t \sum_{i=1}^{n} P_i(\vartheta) T^0_{ij}(t - \vartheta, \vartheta)\, d\vartheta. \tag{1}$$

The total magnitude of $E_j(t)$ may be derived in a similar way from T_{ij}, yielding

$$E_j(t) = \int_0^t \sum_{i=1}^{n} P_i(\vartheta) T_{ij}(t-\vartheta, \vartheta) \, d\vartheta. \tag{2}$$

The portion of this due to previous conditioning is then

$$Q_j(t) = E_j(t) - S_j(t). \tag{3}$$

To derive an expression for the variation of $T_{ij}(\delta, t)$ with time, we may employ the following method. Suppose we have two time intervals, during the first of which, $\xi \to \xi + d\xi$, type i stimulation is being applied, and a later interval, $\eta \to \eta + d\eta$, during which there is both a positive conditioned and an unconditioned tendency to make a type j response. By the fundamental fact of conditioning, if conditioning has not already reached the saturation point, this will result in an increased tendency for future occurrences of type i stimuli to evoke type j responses after a time lag of magnitude $\eta - \xi$—that is to say, an increment of $T_{ij}(\delta, t)$ for $t \geqq \eta$ and $\delta = \eta - \xi$ will be set up. Let us denote this increment by $\Delta_{\xi\eta} T_{ij}(\delta, t) \, d\xi \, d\eta$, supposing that to a sufficient approximation this quantity is constant for all points ξ and η in $d\xi$ and $d\eta$ respectively. Since we shall ultimately allow these intervals to approach zero as a limit, this constitutes no real loss in generality.

First, it is a well-established experimental fact that

$$\Delta_{\xi\eta} T_{ij}(\eta - \xi, t)$$

decreases with an increase in the separation $\eta - \xi$ of the stimulus and the response-tendency. This has been observed directly, in the form of rapidly augmenting difficulty of producing trace-conditioning with greater distance, by Kappauf and Schlosberg (32), H. M. Wolfle (54), Rodnick, (43), Switzer (50), and Pavlov (37). In some of these studies, it is true, it is found that there is an initial increase in facility of conditioning with distance between the conditioned and the unconditioned stimuli, but this may be explained on the hypothesis of a diminished contribution to conditioning from temporal generalization, which as an effect is a consequence of the theory. Pavlov (38, 39) supposes that this gradient with trace-conditioning is due to the fact that the reaction-tendency is here conditioned, not to the stimulus itself, but to the arrival of a certain stage in the decay of a condition set up in the subject by the stimulus, and called its *stimulus-trace*; apart from the obscurity of the notion itself, (which has been used extensively by Hull, especially in 27), it is sometimes assumed that

this trace has a numerical magnitude, which declines with time at a rate proportional to its own magnitude, or according to a negative exponential law; and then that the increment in a trace-conditioning resulting from a given reinforcement is proportional to the strength at the time of the response of the trace originating in the stimulus—an assumption which seems to be in gross agreement with the data cited. We shall consequently make this supposition in its quantitative form, although without making any ontological assumptions as to the nature of the "stimulus trace" which furnishes its rationale. We shall suppose that $\Delta_{\xi\eta} T_{ij}(\eta - \xi, \eta)$ is proportional (1), to $P_i(\xi) e^{-\alpha(\eta-\xi)}$, for a suitable constant α—this is the strength at $d\eta$ of the "trace" from $d\xi$— and (2), to $S_j(\eta) - Q_j(\eta)$, or the amount whereby the unconditioned part of the type j response-tendency exceeds the conditioned part, both evaluated at η. This last assumption is in agreement with the studies of Campbell and Hilgard (7), who found a correlation of .55 ± .06 between strengths of conditioned and unconditioned responses; of Schlosberg (45), and Campbell (6), who found less significant correlations, as to the use of $S_j(\eta)$: for the subtraction of $Q_j(\eta)$ we have considerable evidence which shows that the increment in conditioning which results from a single reinforcement grows less and less as conditioning proceeds, and finally ceases—i.e., the conditioning curve is negatively accelerated. Among the curves of this sort we may mention those of Hovland (25) especially, together with others reproduced in (18). The S-shape of some conditioning curves will be discussed in a later study, together with its possible relation to temporal generalization. If we choose units for P_i properly, our assumptions yield

$$\begin{aligned}\Delta_{\xi\eta} T_{ij}(\eta - \xi, \eta)\, d\xi d\eta \\ = P_i(\xi)\,[S_j(\eta) - Q_j(\eta)]\, e^{-\alpha(\eta-\xi)}\, d\xi d\eta \,.\end{aligned} \tag{4}$$

Now suppose $\delta \neq \eta - \xi$. We may expect that $\Delta_{\xi\eta} T_{ij}(\delta, \eta)$ will be smaller than before, but scarcely that it vanishes; the reinforced interval between the conditioned stimulus and the eduction of its response is very far from being so rigidly maintained in experimental trials (See E. J. Rodnick, 43). Perhaps the principal source of this lack of constancy may be sought in more or less accidental physiological fluctuations. If we make the usual assumption, that these fluctuations are large in number and independent, and hence are distributed in an approximation to normality about η, say with a precision β, we shall obtain

$$\Delta_{\xi\eta}\, T_{ij}(\delta\,,\eta)\,d\xi d\eta = P_i(\xi)\,[S_j(\eta) - Q_j(\eta)]\; e^{-\alpha\delta-\beta^2(\eta-\xi-\delta)^2}\,d\xi d\eta\,, \tag{5}$$

where we of course now evaluate the "stimulus trace" at $\xi + \delta$ instead of at η as before. It may be remarked that on account of the presence of the "stimulus trace" from $\xi \rightarrow \xi + d\xi$ in (5), in the factor $e^{-\alpha\delta}$, the actually observed variations in the interval between a stimulus-presentation and the evocation of its conditioned response will be skewed positively, perhaps even settling about a new mean, much closer to ξ than is η. This is the expression of a fact well-known experimentally (44, 50, and 45), a circumstance which lends additional plausibility to our choice of assumptions.

Finally, this increment in $T_{ij}(\delta\,,\,t)$, set up for $t = \eta$, may be expected to decay with time, from the accumulation of adventitious retroactive inhibition (whose effect, incidentally, our theory will enable us to predict, if it be specified in detail, in accord with the hypothesis which explains such inhibition in terms of counter-conditioning; see section 6) or from some other cause, in such a way as to result in its slow but eventually almost complete loss. Here again, there is evidence which leads us to suppose an exponential law of decay (20, and the work of Ebbinghaus), and in doing so, we shall be in harmony with previous theories (Hull, 27, 28). Combining this with (5), we shall finally obtain

$$\Delta_{\xi\eta}\, T_{ij}(\delta\,,t)\,d\xi d\eta = P_i(\xi)\,[S_j(\eta) - Q_j(\eta)]\; e^{-\alpha\delta-\beta^2(\eta-\xi-\delta)^2-\gamma(t-\eta)}\,d\xi d\eta\,, \tag{6}$$

for a suitable decay-constant γ.

Now to obtain the total value of $T_{ij}(\delta\,,\,t)$ at any time, we must consider (1), the unconditioned part $T^0_{ij}(\delta\,,\,t)$, and (2), contributions of magnitude (6) from all pairs of intervals $d\xi\,,\,d\eta\,,\,d\xi$ preceding $d\eta$, which occur before t. If these intervals are still considered as finite in extension, this result will be only approximate; if, however, we allow them to approach zero in extension, while increasing indefinitely in number, so that they still cover the whole region $0 \rightarrow t$, then, taking the limit, our result will become exact; and we shall have

$$T_{ij}(\delta\,,t) = T^0_{ij}(\delta\,,t) + \int_0^t d\eta \int_0^\eta P_i(\xi)\,[S_j(\eta) - Q_j(\eta)]\; \mathrm{Exp}\,\{-\,\alpha\,\delta - \beta^2(\eta - \xi - \delta)^2 - \gamma(t-\eta)\}\,d\xi\,. \tag{7}$$

This expression may be substituted in (2); taking account of (1) and

(3), and changing the order of integration by Dirichlet's rule, we shall find

$$Q_j(t) = \int_0^t S_j(\eta) \int_\eta^t d\vartheta \int_0^\eta F(\vartheta, \xi) \operatorname{Exp} \{ -\alpha(t-\vartheta) - \beta^2(\eta - \xi - t$$

$$+ \vartheta)^2 - \gamma(\vartheta - \eta) \} \, d\xi - \int_0^t Q_j(\eta) \int_\eta^t d\vartheta \int_0^\eta F(\vartheta, \xi) \tag{8}$$

$$\operatorname{Exp} \{ -\alpha(t-\vartheta) - \beta^2(\eta - \xi - t + \vartheta)^2 - \gamma(\vartheta - \eta) \} \, d\xi ,$$

where we have put

$$F(\vartheta, \xi) = \sum_{i=0}^{N} P_i(\xi) \, P_i(\vartheta) ,$$

and, of course, $\operatorname{Exp}(x)$ is a more convenient form of e^x. Upon abbreviating

$$\left\| \begin{matrix} \alpha, \beta, \gamma \\ \eta, t \end{matrix} \right\| = \int_\eta^t d\vartheta \int F(\vartheta, \xi) \operatorname{Exp} [-\alpha(t-\vartheta) - \beta^2(\eta - \xi - t + \vartheta)^2$$

$$- \gamma(\vartheta - \eta)] \, d\xi , \tag{9}$$

this may be written

$$Q_j(t) = \int_0^t \left\| \begin{matrix} \alpha, \beta, \gamma \\ \eta, t \end{matrix} \right\| S_j(\eta) \, d\eta - \int_0^t \left\| \begin{matrix} \alpha, \beta, \gamma \\ \eta, t \end{matrix} \right\| \times$$

$$Q_j(\eta) \, d\eta . \tag{10}$$

(10) is a Volterra integral equation of the second kind for the determination of $Q_j(t)$, with a kernel

$$K(t, \eta) = - \left\| \begin{matrix} \alpha, \beta, \gamma \\ \eta, t \end{matrix} \right\| .$$

Since all other quantities occurring in (8) are known from the experimental situation, if we can solve (8) for $Q_j(t)$, this will constitute a solution to our problem. That this can be done follows from the facts, (1) that $F(\vartheta, \xi) = \sum_i P_i(\xi) \, P_i(\vartheta)$ may have at most a finite number of regularly-distributed discontinuities, owing to the fact that the presentation of some stimulus may have begun or ended abruptly; hence $\left\| \begin{matrix} \alpha, \beta, \gamma \\ \eta, t \end{matrix} \right\|$ is continuous; and (2), $S_j(\eta)$ is also continuous by (1), for any physically possible form of T^0_{ij}. We shall accordingly find it possible to find Q_j, and hence the total reaction tendency $E_j = Q_j + S_j$, in terms of $S_j(t)$, the unconditioned tendency to make a type j response, and the functions P_i, which measure the stimulation occurring up to t in the experimental situation; and both of the

functions may be determined by observation, the latter directly, the former through equation (1). The explicit solution of (10) may be made fairly easily in series. Suppose we define $\left\| \begin{matrix} \alpha,\beta,\gamma \\ \eta,t \end{matrix} \right\|^{(n)}$, where n is a positive integer, by the following recursion:

$$\left\| \begin{matrix} \alpha,\beta,\gamma \\ \eta,t \end{matrix} \right\|^{(1)} = - \left\| \begin{matrix} \alpha,\beta,\gamma \\ \eta,t \end{matrix} \right\|,$$

$$\left\| \begin{matrix} \alpha,\beta,\gamma \\ \eta,t \end{matrix} \right\|^{(n+1)} = \int_\eta^t - \left\| \begin{matrix} \alpha,\beta,\gamma \\ t,\zeta \end{matrix} \right\| \cdot \left\| \begin{matrix} \alpha,\beta,\gamma \\ \zeta,t \end{matrix} \right\|^{(n)} d\zeta .$$

Now define the function $\Psi(n, t)$ called the *solvent kernel* of $-\left\| \begin{matrix} \alpha,\beta,\gamma \\ \eta,t \end{matrix} \right\|$, by the infinite series

$$\Psi(\eta,t) = \sum_{n=1}^{\infty} \left\| \begin{matrix} \alpha,\beta,\gamma \\ \eta,t \end{matrix} \right\|^{(n)}, \tag{11}$$

which converges throughout the finite range. Then the solution of (10) may be written

$$\begin{aligned} Q_j(t) = \int_0^t \left\| \begin{matrix} \alpha,\beta,\gamma \\ \eta,t \end{matrix} \right\| S_j(\eta)\, d\eta + \int_0^t \Psi(\eta,t)\, d\eta \times \\ \int_0^\eta \left\| \begin{matrix} \alpha,\beta,\gamma \\ \zeta,\eta \end{matrix} \right\| S_j(\zeta)\, d\zeta , \end{aligned} \tag{12}$$

as may easily be verified by substituting the series (11) in (12) and the resulting value of Q_j in (10). This expression may be simplified somewhat: remembering that

$$\int_\eta^t \left\| \begin{matrix} \alpha,\beta,\gamma \\ t,\zeta \end{matrix} \right\| \Psi(\zeta,\eta) d\zeta = \Psi(\eta,t) - \left\| \begin{matrix} \alpha,\beta,\gamma \\ \eta,t \end{matrix} \right\|,$$

which is an immediate consequence of (12), we shall have, applying Dirichlets' rule and reliterating,

$$\begin{aligned} Q_j(t) &= 2 \int_t^0 \left\| \begin{matrix} \alpha,\beta,\gamma \\ \eta,t \end{matrix} \right\| S_j(\eta)\, d\eta - \int_0^t \Psi(\eta,t) S_j(\eta) d\eta \\ &= \int_0^t \left\{ 2 \left\| \begin{matrix} \alpha,\beta,\gamma \\ \eta,t \end{matrix} \right\| - \Psi(\eta,t) \right\} S_j(\eta) d\eta . \end{aligned} \tag{13}$$

This solution, as mentioned before, is entirely general, subject of course to the assumptions we have made, and the independence requirements of the first three sections.

3. Extinction and the reaction threshold.

We shall partially follow Razran (42) and others in the present discussion, in attributing the effects of experimental extinction and allied forms of inhibition to the combined action of three groups of forces. The first of these, which involves setting up a counter-conditioning between the conditioned stimulus and some extraneous and more or less adventitious reaction, is covered by the considerations developed in the previous section, except as regards its interference with the original conditioned reaction; and this will be discussed in section 6. The second sort of influence, an "inverse conditioning" between the stimulus and the same reaction, is also a consequence of our previous discussion, as may be seen at once from (6). Here, under the circumstances of an extinction trial, in which the unconditioned stimulus is not presented, $S_j(\eta)$ vanishes; but on account of previous conditioning and the presentation of the conditioned stimulus, $Q_j(\eta) > 0$. Accordingly, (6) becomes negative, and the strength of conditioning suffers a decrement, proportional in amount to its own magnitude, rather than an increase as before.

The two components of extinction that we have mentioned are essentially like conditioning itself in the rapidity of their acquisition, rate of decay, and other properties. It is notorious, however, that most of the known properties of experimental extinction and allied inhibitions are as different from those of any form of conditioning as it is possible to be: extinction is obtained very much more rapidly than conditioning (Hovland, 25); it decays, to a partial extent, very quickly, a phenomenon known as "spontaneous recovery" (See Hovland, **25**; Pavlov, **37**, **38**; Finch and Culler, **13**; D. G. Ellson, **12**; C. H. Coombs, **10**; E. R. Hilgard and D. G. Marquis, **19**, **21**; for further references also Hilgard and Marquis, **18**; and Razran, **43**, the latter being an excellent summary of the facts regarding extinction); the drugs which accelerate it usually retard conditioning, and conversely; (see Hilgard and Marquis, 18, for a discussion of this with references) it can be destroyed temporarily by any sudden stimulus ("disinhibition"), a procedure which has no effect on conditioning at all; and, in general, all the influences which operate to favor extinction usually reduce conditioning; and conversely, agencies which aid conditioning tend to diminish the facility wherewith extinction is obtained. This last is evidenced by the correlations between extinction and conditioning which have been computed from experimental populations (summarized in **18**, p. 119), and are predominantly high and negative.

This evidence tends to suggest very strongly that there is a component of experimental extinction which is quite different in its law

of variation from ordinary conditioning: in our present discussion we shall take this into account by introducing a *variable reaction-threshold,* itself subject to conditioning, but of the relatively temporary and quick-decaying kind which is characteristic of the inhibitory phenomena. In previous macroscopic theories (e.g. Hull, 27), the reaction-threshold has been taken as constant or at most subject to random variations; but there is no necessity for this, and any apparent one probably results from a confusion between the "reaction-threshold" as we understand it, having reference to gross behavioral phenomena, and the "reaction-threshold" of neurology, which does not stand to the former in any simple relation. The latter is perhaps a physiological constant; but the former may be subjected to variation of any convenient kind. Indeed, a good quantity of evidence appears to suggest the hypothesis that experimental extinction and related inhibitions are due in large part to a temporary rise in the reaction-threshold, understanding this latter as is usual in psychology: as examples of this, of a qualitative sort, we may mention the following.

First, it is well known (25, 44, 18, p. 127) that after a conditioning has been completely extinguished, a sudden fear-evoking stimulus will effect a temporary resurgence of the conditioning; this phenomenon is generally called "disinhibition." If we suppose that a startling stimulus causes a suddenly heightened state of excitation in the organism, which may be interpreted as a temporary lowering of all reaction-thresholds, our hypothesis would lead us to expect just such a loss of inhibition.

Again, it is known that depressant drugs, such as bromides, retard conditioning while expediting extinction, whereas stimulants, such as benzedrine and caffeine, conversely accelerate conditioning and retard extinction. Here also, if we make the natural supposition that depressants raise all the reaction-thresholds, while excitants reduce them, then our hypothesis, which asserts that extinction is a conditioned rise in the threshold, would predict more rapid extinction in the former case, since less further net increase would be required to inhibit the reaction; but slower conditioning, since more reinforcement would be needed to enable the response-tendency to attain the augmented threshold. As remarked, this is just what occurs. A similar account can be given of many cases involving phenomena such as "sensitization," or "pseudoconditioning" (18, chap. I).

A more precise statement can be made readily, by proceeding in relation to the considerations of the previous section. Let us denote the reaction-threshold for type j responses by $R_j(t)$, and its initial,

"normal" value by R_{0j}. Now, if we remember that experimental extinction appears to occur wherever there is conditioned reaction-tendency, whether there is unconditioned response-tendency there also or not, and hence is not likely to depend on S_j, and also that it appears to increase and decrease more rapidly than ordinary conditioning, the same reasoning which led to (6) will yield

$$\Delta_{\xi\eta} T^R{}_{ij}(\delta, t) = \nu P_i(\xi) Q_j(\eta)\, e^{-a\delta-\beta'^2(\eta-\xi-\delta)^2-\gamma'(t-\eta)}, \tag{13'}$$

where the expression on the left denotes the increment in the strength of conditioning between P_i and R_j, of which $T^R{}_{ij}(\delta, t)$ is the total value. Precisely as before, we may integrate (13′) to obtain

$$T^R{}_{ij}(\delta, t) = T^{R0}{}_{ij}(\delta, t) + \nu \int_0^t Q_j(\eta)\, d\eta \int_0^\eta P_i(\xi) \times e^{-a\delta-\beta'^2(\eta-\xi-\delta)^2-\gamma'(t-\eta)}\, d\xi, \tag{14}$$

and substituting this in the analogue of (2) for R_j and T^R, and abbreviating by (9), we derive

$$R_j(t) = R_0 + \nu \int_0^t \left\| \begin{matrix} \alpha, \beta', \gamma' \\ \eta, t \end{matrix} \right\| Q_j(\eta)\, d\eta, \tag{15}$$

which specifies R_j directly in terms of the conditioning process. Of the constants α, β', γ', γ in (15), α remains the same as in 9, since it represents the strength of the "stimulus trace" from ξ at η, which is the same in the two cases; the relation of β' to β is unknown, although the general liability of R_j would lead us to suppose $\beta' > \beta$; and the very much more rapid decay of experimental extinction as compared with ordinary conditioning implies $\gamma' > \gamma$. We have mentioned before that probably $\nu > 1$. Incidentally, the exponential law for spontaneous recovery, involved in the term $e^{-\gamma(t-\eta)}$ in (13′), is well attested by recent experiments: among these we may mention especially those of D. G. Ellson (12) and Hovland (25).

The hypothesis (15) is sufficient to explain, qualitatively at least, all the principal types of inhibition. First among these we may take the so-called "inhibition of delay," an observed effect in which, after trace-conditioning has taken place between a stimulus i and a reaction j, there is an active inhibition of the response between the presentation of i and the elicitation of j, so that stimuli which even unconditionally evoked j will fail to do so, or the strength of the response will undergo marked coarctation (see e. g. Pavlov, 37, and Rodnick, 44). For, suppose that in the routine of conditioning, i is ordinarily presented at a time, say t, and j appears, due to unconditioned stimulation, at $t + \delta$; but let it be the case, as in experimental trace-condi-

tioning routines, that while $S_j(t + \delta) > R_j(t + \delta) > 0$, $S_j(t + \varepsilon)$, for $0 < \varepsilon < \delta$, is negligible. Nevertheless, a tendency to respond, $E_j(t + \varepsilon) = Q_j(t + \varepsilon)$, which is wholly the result of conditioning, will be built up at that point, owing to the generalization from $t + \delta$, mentioned in the derivation of (3) above. Now when this has occurred, the situation at $t + \varepsilon$ is exactly the same as in experimental extinction; by (15) the reaction-threshold R_j will be conditioned to increase at that point, and the net result will be a strong inhibition. As remarked, this is what occurs. This explanation in terms of our hypothesis, incidentally, has two further consequences of interest. First, the inhibition of delay will be subject to spontaneous recovery; this is remarked upon by Switzer (50), and Rodnick (44). Secondly, after the inhibition of delay has disappeared by spontaneous recovery, the accumulated Q_j at $t + \varepsilon$ will be active reaction-tendency, and if great enough will tend to shorten the reaction-latency or even elicit a second anticipatory reaction before the primarily conditioned one. Effects of this sort have been observed by Rodnick (43), in considerable detail.

A somewhat simpler application of (15) may be seen in the so-called "inhibition of reinforcement." A glance at (15) makes evident the fact that, as a series of reinforcements is traversed, Q_j grows larger and larger, and consequently R_j begins to increase also, the more rapidly the further conditioning proceeds. When Q_j reaches the saturation point, and no longer rises, the increment in R_j continues apace, and the net result is an apparent *reduction* in the strength of conditioning. This is especially marked when the later conditioning reinforcements are massed together in large numbers, so that R_j may increase faster than it can be lost by spontaneous recovery. The diminution from this source would be expected to disappear after a relatively short interval. This phenomenon does in fact occur, and has all the properties we have mentioned: it is observed by Hovland (26), and is discussed by Razran (42), and by Hilgard and Marquis (18, p. 124).

Another sort of phenomenon is that sometimes called "conditioned inhibition." In this effect, a given stimulus i is conditioned to a reaction j, and then paired (successively) with another stimulus k. i and k are then presented together, without the unconditioned stimulus, until the reaction j is extinguished. Then, by our hypothesis, i and k are both conditioned to evoke a rise in R_j. With respect to i, a few reinforcements are sufficient to remove its extinction, so that it again elicits j. When paired with k, however, to which an increment in R_j is still conditioned, we should predict that, under suit-

able conditions, the response due to i would be partially or completely inhibited. In this case we may call k a *conditioned inhibitor* of the reaction j. This part of conditioned inhibition is of course quite transitory; by more extended efforts, however, a more permanent variety can be established, according to the mechanisms of counter-conditioning and inverse conditioning remarked upon at the beginning of the present section. These expectations are all verified observationally; for an experimental study see (18) p. 126. It is upon these more permanent effects also that we rely to account for the significant part of experimental extinction which is relatively permanent in nature, and for the inculcation of discrimination by the "method of contrasts," which Hovland found (24, 26) to consist of two parts, one analogous to ordinary adaptation in its rapid spontaneous loss and susceptibility to disinhibition, the other being the relatively permanent inhibition found in previous studies (e.g., Hilgard and Humphries, 20), as would be required by a hypothesis such as ours.

The law of variation of the reaction-threshold which we have embodied in (15), however, must not be applied indiscriminately. There are a number of experimental indications which appear to suggest that with respect to many human reactions, which we should ordinarily say are to some extent under "voluntary control," the reaction-threshold, even more than the other variables of conditioning, is determined by factors of this sort, and is consequently not subject to conditioning, at any rate to the extent of being in large measure controlled by such processes. By this, of course, we mean that the threshold depends partially upon the exigencies of other influences involving the subject than are contained within our present theoretical framework—e.g., by the verbal instructions issued to the subject at the beginning of the experiment. Effects of this kind would tend to explain many empirical findings: for example, that experimental extinction fails to occur in conditioning situations with a frequency which appears to vary directly with the extent to which we should say that the function in question is "controlled voluntarily" (18). Again, we may adduce the experiment of J. Miller (35), who considered the effect on conditioning of the eyelid response when subjects were issued instructions at the beginning of the experiment (a) to inhibit actively, and (b) to refrain from control of the response entirely. In the group which was issued the instructions (a), conditioning was slower and extinction quicker than in the control group; whereas the instructions (b) had precisely the converse effect. This is what we should expect on the hypothesis of a raised reaction-threshold in the first case and a diminished one in the other. Further, in cases of se-

rial learning of nonsense syllables, to reproduce each rote-syllable immediately upon presentation of its predecessor in the series, the eduction of reactions in general appears to be syntonic with the beginning of a syllable-presentation, even though the reaction-tendency, as determined from (12), may be greater at another point during the same presentation (27)—i.e., instructions are followed by modifying the reaction-threshold. In general, we may say that the major difference in experimental behavior between the conditioned response and human learning of non-symbolic material resides in the almost total absence in the latter of all the complicated forms of inhibition which are characteristic of conditioning in very many of its aspects—the same procedure which totally extinguishes conditioning will have no effect in a learning experiment—and the assumption of voluntary control of the reaction-threshold in the one case, and accordance with (15) in the other, would appear to reconcile this difference excellently. This difference in observational properties, it may be remarked, has proved a strong objection to the previous behavioristic theories which have attempted to find in learning only a more complex and multiform kind of conditioning. This hypothesis, of course, does not remove the reactions considered, which constitute nearly all of human learning, from the scope of our considerations: in the first place, the threshold cannot be determined wholly at will, and in cases of learning with the subjects' cooperation, where it is perhaps forced down to its minimum possible level, — except for a few regular differences of the kind appearing in the case of serial rote-learning—the course of learning will still be determined as in section 1. The only difference in fact, is that here the subject may refuse to cooperate, and fail voluntarily to learn, whereas in the conditioning of involuntary and autonomic functions this is not possible. Finally, if a criterion for the "voluntary" nature of a reaction is sought, we may perhaps reply in terms of (15): a response-function is involuntary to an extent depending directly upon the presence and regularity of experimental extinction.

References

1. Bass, N. J. and C. L. Hull. The irradiation of a tactile conditioned reflex in man. *J. comp. Psychol.*, 1934, 17, 47-65.

2. Brogden, W. C. and E. Culler. Experimental extinction of higher order responses. *Amer. J. Psychol.*, 1935, 47, 663-669.

3. Brogden, W. C., E, A, Lippman, and E. Culler. The role of incentive in conditioning and extinction. *Amer. J. Psychol.*, 1938, 51, 107-117.

4. Brunswik, E. S. Probability as a determiner of rat behavior. *J. exp. Psychol.*, 1939, 75, 175-197.

5. Bugelski, R. Extinction with and without sub-goal reinforcement. *J. exp. Psychol.*, 1940, 26, 121-133.

6. Campbell, A. A. The interrelations of two measures of conditioning in man. *J. exp. Psychol.*, 1938, 22, 244-268.

7. Campbell, A. A. and E. R. Hilgard. Individual differences in ease of conditioning. *J. exp. Psychol.*, 1931, 14, 561-571.

8. Carnap, R. Testability and meaning. *Phil. of Sci.*, 1935 & 1936, 3, 419-471, and 4, 1-40.

9. Carnap, R. Logical foundations of the unity of science. *Int. Encyc. of Unified Sci.*, 1939, vol. 1, no. 1, 42-62.

10. Coombs, C. H. Adaptation of the galvanic skin response to auditory stimuli. *J. exp. Psychol.*, 1938, 22, 244-268.

11. Courant, R. and D. Hilbert. Methoden der Mathematischen Physik. Bd. 1. Berlin: Julius Springer, 1931.

12. Ellson, D. G. Spontaneous recovery of the galvanic skin response as a function of the recovery interval. *J. exp. Psychol.*, 1939, 25, 586-600.

13. Finch, G. and E. Culler. The relation of forgetting to experimental extinction. *Amer. J. Psychol.*, 1933, 47, 656-662.

14. Gullikson, H. Studies of transfer of responses: I. Relative versus absolute factors in the discrimination of size by the white rat. *J. genet. Psychol.*, 1932, 40, 37-51.

15. Gullikson, H. A rational equation of the learning curve based on Thorndike's law of effect. *J. gen. Psychol.*, 1934, 11, 395-439.

16. Gullikson, H. and D. L. Wolfle. A theory of learning and transfer. *Psychometrika*, 1938, 3, 127-449 and 225-251.

17. Hilgard, E. R. An algebraic analysis of conditioned discrimination in man. *Psychol. Rev.*, 1938, 45, 472-496.

18. Hilgard, E. R. and D. G. Marquis. Conditioning and learning. New York: D. Appleton-Century Company, 1940.

19. Hilgard, E. R. and D. G. Marquis. The course of acquisition and retention of conditioned eyelid responses in man. *J. exp. Psychol.*, 1936, 19, 227-247.

20. Hilgard, E. R. and L. G. Humphries. The retention of conditioned discrimination in man. *J. gen. Psychol.*, 1938, 19, 111-125.

21. Hilgard, E. R. and D. G. Marquis. Acquisition, extinction, and retention of conditioned eyelid responses to light in dogs. *J. comp. Psychol.*, 1935, 19, 29-38.

22. Horst, A. P. A new method for determining the absolute affective value of a series of stimulus situations. Thesis, University of Chicago, 1931; published in part in *J. educ. Psychol.*, 23, no. 6, Sept. 1932.

23. Hovland, C. I. The generalization of conditioned responses: I. The sensory generalization of conditioned responses with ranging frequencies of tone. *J. gen. Psychol.*, 1937, 17, 125-148.

24. Hovland, C. I. The generalization of conditioned responses: II. The gen-

eralization of conditioned responses with varying intensities of tone. *J. genet. Psychol.*, 1937, **30**, 279-291.

25. Hovland, C. I. The generalization of conditioned responses: III. Extinction, spontaneous recovery, and disinhibition of conditioned and of generalized responses. *J. exp. Psychol.*, 1937, **21**, 47-61.

26. Hovland, C. I. The generalization of conditioned responses: IV. The effect of varying amounts of reinforcement upon the degree of generalization of conditioned responses. *J. exp. Psychol.*, 1937, **21**, 261-276.

27. Hull, C. L. et al. Mathematico-deductive theory of rote-learning. New Haven: Yale University Press, 1940.

28. Hull, C. L. The problem of stimulus-equivalence in behavior theory. *Psychol. Rev.*, 1939, **46**, 9-30.

29. Hull, C. L. Mind, mechanism, and adaptive behavior. *Psychol. Rev.*, 1937, **44**, 1-32.

30. Humphries, L. H. The effect of random alternations of reinforcement and extinction on the acquisition and extinction of conditioned eyelid responses. *J. exp. Psychol.*, 1939, **25**, 141-157.

31. Humphrey, G. The nature of learning in its relation to the living system. New York: D. Appleton Century, 1938.

32. Kappauf, W. S. and H. Schlosberg. Conditioned responses in the white rat: III. Conditioning as a function of the length of the period of delay. *J. genet. Psychol.*, 1937, **50**, 27-45.

33. Kellogg, W. N. and I. S. Wolf. The nature of the reaction retained after several varieties of conditioning in the same subjects. *J. exp. Psychol.*, 1939, **24**, 366-383.

34. Kupalov, P. S. and W. H. Gantt. The relation between the strength of the conditioned stimulus and the size of the resulting conditioned reflex. *Brain*, 1927, **54**, 83-98.

35. Miller, J. The effect of inhibitory and facilitory attitudes on eyelid conditioning. Ph.D. Dissertation. Yale University. Abstract in *Psychol. Bull.*, 1939, **36**, 377-378.

36 Münzinger, K. F. and C. C. Dove. Serial learning: I. Gradients of uniformity and variability produced by success and failure. *J. gen. Psychol.*, 1937, **16**, 403-413.

37. Pavlov, I. P. Conditioned reflexes: an investigation of the physiological activity of the cerebral cortex. London: Oxford University Press, 1927.

38. Pavlov, I. P. Lectures on conditioned reflexes. (trans. and ed. by W. H. Gantt) New York: International Publishers, 1928.

39. Rashevsky, N. Mathematical biophysics, physicomathematical foundations of biology. Chicago: University of Chicago Press, 1938.

40. Rashevsky, N. Advances and applications of mathematical biology. Chicago: University of Chicago Press, 1940.

41. Razran, G. H. S. Transposition of relational responses and generalization of conditioned responses. *Psychol. Rev.*, 1938, **45**, 532-537.

42. Razran, G. H. S. The nature of the extinctive process. *Psychol. Rev.*, 1939, **46**, 264-297.

43. Rodnick, E. H. Characteristics of delayed and trace-conditioned responses. *J. exp. Psychol.*, 1937, 20, 409-425.

44. Rodnick, E. H. Does the interval of delay of conditioned responses possess inhibitory properties? *J. exp. Psychol.*, 1937, 20, 507-527.

45. Schlosberg, H. An investigation of certain factors related to ease of conditioning. *J. gen. Psychol.*, 1932, 7, 328-342.

46. Skinner, B. F. The behavior of organisms: an experimental analysis. New York: D. Appleton-Century, 1938.

47. Spence, K. W. Analysis of the formation of visual discrimination habits in the chimpanzee. *J. comp. Psychol.*, 1937, 22, 77-111.

48. Spence, K. W. The differential response in animals to stimuli ranging within a single dimension. *Psychol. Rev.*, 1937, 44, 430-444.

49. Switzer, S. A. Anticipatory and inhibitory characteristics of delayed conditioned reactions. *J. exp. Psychol.*, 1937, 20, 409-425.

50. Switzer, S. A. Anticipatory and inhibitory characteristics of delayed conditioned reactions. *J. exp. Psychol.*, 1934, 17, 603-620.

51. Thorndike, E. L. The psychology of wants, interests, and attitudes. New York: D. Appleton-Century, 1935.

52. Volterra, V. Theory of functionals. London: Blackie and Son, Limited, 1931. Trans. by Miss M. Long.

53. Williams, S. B. Resistance to extinction as a function of the number of reinforcements. *J. exp. Psychol.*, 1938, 23, 506-522.

54. Wolfle, H. M. Conditioning as a function of the interval between the conditioned and the original stimulus. *J. gen. Psychol.*, 1932, 7, 80-103.

55. Youtz, R. E. P. Reinforcement, extinction, and spontaneous recovery in a non-Pavlovian reaction. *J. exp. Psychol.*, 1938, 22, 305-318.

56. Youtz, R. E. P. The change with time of a Thorndikian response in the rat. *J. exp. Psychol.*, 1938, 23, 128-134.

57. Youtz, R.E.P. The weakening of one Thorndikian response following the extinction of another. *J. exp. Psychol.*, 1939, 24, 294-304.

PSYCHOMETRIKA—VOL. 8, NO. 1
MARCH, 1943

FACTORIAL EQUATIONS FOR TESTS OF ATTENTION*

JOHN RICHARD WITTENBORN
YALE UNIVERSITY

Tests designed to measure what was conceived to be attention were found by factor analysis to involve a factor which is independent of the factors of rote memory, visual-space, number, or perception. To a large extent, at least, this attention factor is independent of content and of mode of presentation of test material. The tests in which the variance is mainly dependent upon this factor are those involving a high degree of sustained or relatively continuous mental effort.

A doctoral dissertation submitted by Brother Rogation Philip to the Faculty of Philosophy of the Catholic University of America in 1928 (4) describes what is probably the first attempt to measure attention by a battery of tests. Brother Philip proposed to construct a battery of pencil-and-paper tests which would yield a high multiple correlation with a tedious laboratory measurement of attention devised by Woodrow (10). Woodrow's criterion of attention was the inverse of the difference between reaction time with widely varying preparatory intervals and reaction time with a uniform preparatory interval of 2 seconds. Brother Philip's battery of five tests correlated .596 with this criterion.

Evidence indicating the existence of a factor named attention has been found in several factor analyses of the intercorrelations of test scores. The *Preliminary Report on Spearman-Holzinger Unitary Trait Study,* which appeared in 1935 (2), showed that tests devised by Brother Philip determine one of the factors appearing in the Hollow Staircase factor pattern.

In an analysis of 52 mental tests published by Woodrow in 1939 (11), two of the tests devised by Brother Philip were among the tests highly saturated with a new factor. The name attention, tentatively

* An abstract of a thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Psychology in the Graduate School of the University of Illinois, 1942.

The writer wishes to express his gratitude to Professor Herbert Woodrow who suggested the present problem and under whose guidance the investigation was conducted. A further debt of gratitude is acknowledged to the enlisted men at the Air Corps Technical School at Chanute Field, Illinois, who served as subjects and to Capt. T. W. Harrell and Capt. Richard Faubion.

applied to this factor, was suggested by the kind of operations in which the factor was manifested. Two of the tests determining this new factor, Philip's Alphabet and Following Directions from Otis (3), are included in the present study.

Woodrow's analysis of factors in improvement with practice in 1939 (12) showed a factor similar to the one described as attention in another of his studies (11). This factor also showed a high correlation with Philip's Alphabet and Otis' Following Directions.

In order to investigate and to define more clearly the ability of attention suggested by the above studies, the present study was undertaken. The plan of the investigation involved two chief steps:

1. Devising new tests of attention.
2. Analyzing by factor analysis the intercorrelations of a battery of tests which included, in addition to the newly devised tests, tests of certain well established factors.

The primary purpose of the investigation was to determine the factorial composition of the tests devised as tests of attention.

In the construction of the tests a large number of items or performances were devised. These items were regarded as equivalent or similar to the performances distinguishing or characterizing those tests which contributed to the factors tentatively identified as attention factors by previous investigators (2, 11, 12). Particular consideration was given to the tests of Philip used by Holzinger (2) and Woodrow (11, 12), as well as to certain other tests used by Woodrow, termed Vertical Copying (11) and Directions (11, 12). In the selecting or devising of the items for new tests, certain requirements were formulated:

1. The performances should not depend too much upon intellectual level.
2. The tasks should depend to as small a degree as possible upon content and knowledge.
3. The tests should correlate as little as possible with factors heretofore identified.
4. The scores on the tests should depend to a large degree upon a continuous, sustained application of mental effort. The tests should be so constructed that a layman might say they required a high degree of continuous "concentration."

In an attempt to devise a performance that would not depend too much upon a particular kind of knowledge, material which was fa-

miliar to almost all persons was chosen; namely, digits and letters of the alphabet. A variety of tests was tried, and the kind of tests which most dependably required a continuous effort of attention seemed to be those patterned somewhat after tests ordinarily called directions tests; therefore, new tests of this general type were used. To insure a relatively continuous application of attention, the items were presented one at a time at a rate that allowed little opportunity for interpolated activity. The importance of continuous and concentrated work was enhanced by using tasks in which the response to an ensuing item was in part conditioned by responses to previous items. Two possible modes of presentation were considered, visual and auditory. It would be possible to project the items on a screen or to expose them on a revolving drum in a controlled manner; it would also be possible to read the tasks or to present them by means of a phonograph, controlling the time in any desired manner. In the use of either of these methods of presentation, a prepared answer sheet would be a necessity for identifying the responses. After some preliminary investigation, the auditory mode was selected. The determining consideration was mechanical convenience.

The same general procedure was followed in administering each of the tests. The answer sheets were passed out face down, names were written on the back, and at a signal the sheets were turned face up. The examiner read the directions aloud while the subjects read them silently. The sample problems were then read to the subjects, and each answer was explained in terms of the directions. Any answers that were challenged by a subject were explicitly demonstrated. The practice problems were read at a predetermined rate. The sample series were presented at a relatively slow rate, and the practice series were read at a rate which was faster but not so fast as the presentation of the series of the test proper. On a desk before the examiner was a felt-lined box containing a metronome. The metronome, which was set for each change in rate of reading, determined the examiner's rate of presentation. At the end of each practice series, correct answers were given, and any items in question were explained. The rates of presentation and methods of explanation had been experimentally determined and were constant throughout the testing of all groups. Phonographic recordings of the test series, but not of the practice series, were made and employed in the testing of all groups.

Six new tests of attention, numbers 9, 11, 12, 16, 17, and 19, were prepared and phonographically recorded for use in this study. The two tests having the highest correlation with the attention factor were tests 11 and 17. The directions for these two tests, together

with some sample items and a statement of the rates for the presentation of the tasks, follow:

Test 11 Directions

You will hear sets of three digits. Some of the digits will require no response. Some of these sets of digits will be responded to by marking a plus (+) in the answer space reserved for the set. You will mark the set plus under two conditions: 1. when the first number of the set is the largest and the second number of the set is smallest; 2. when the third number of the set is largest and the first number is smallest.

Below are seen sample series of sets with the correct responses indicated. The responses for Series X are corretly indicated, although the numbers comprising the series are not shown.

Samples Series V		Series W		*Practice* Series X	Series Y	Series Z
1 3 4	+	4 1 3	+			
2 7 5		7 3 8		+		
3 8 9	+	9 3 7	+			
9 6 8	+	7 5 2				
8 6 2		9 3 2				
4 1 8		2 7 4		+		
5 3 4	+	3 5 9	+	+		
7 2 6	+	9 5 4				
5 4 8		8 9 7		+		
8 5 3		3 8 2				
7 9 5		6 5 7				
4 7 3		7 4 9				

The sample series V and W were read without the metronome and more slowly than practice series X, Y, and Z, which were read with the metronome set at 72 beats per minute. When the record for the test was cut, the metronome was set at 96 beats per minute. The digits of each set were called on the beat and one beat separated each set. Each of the twelve series of the test proper was announced and six beats separated each series.

Test 17 Directions

Lists of letters selected from the alphabet will be read. You will respond to some of the letters by marking a plus (+). You will respond to other letters by marking a minus (—). Some of the letters will require no response. This answer sheet has been prepared for recording your responses. The spaces below are arranged in lists. The lists below correspond by number to the lists that will be read to you. The letters within each list are numbered, and these numbers correspond to the answer space reserved for that letter.

This is how and when you are to mark the answer spaces: Mark a vowel following a consonant plus (+) and a consonant following a vowel minus (—). When two vowels or consonants occur together, that is, a vowel followed by a vowel or a consonant followed by a consonant, mark the next letter plus (+) no matter what it is. The samples are correctly answered.

Sample:				Practice:			
List X	List x	List Y	List y	List 1	List 2	List 3	List 4
1-A	1	1-X	1	1	1	1	1
2-D	2 —	2-L	2	2	2	2	2
3-C	3	3-O	3 +	3	3	3	3
4-D	4 +	4-P	4 —	4	4	4	4
5-A	5 +	5-R	5	5	5	5	5
6-O	6	6-U	6 +	6	6	6	6
7-L	7 +	7-F	7 —	7	7	7	7
8-M	8	8-C	8	8	8	8	8
9-N	9 +	9-J	9 +	9	9	9	9
10-R	10 +	10-N	10 +	10	10	10	10
11-U	11 +	11-E	11 +	11	11	11	11
12-V	12 —	12-G	12 —	12	12	12	12
13-S	13	13-G	13	13	13	13	13
14-U	14 +	14-D	14 +	14	14	14	14

The sample lists, X and Y, were read without the metronome and more slowly than the practice lists 1, 2, 3, and 4, which were read

with the metronome set at 50 beats per minute. When the record for the test was cut, the metronome was set at 66 beats per minute. Each letter of the lists was called on alternate beats. There was an interval of four beats between the lists during which the next list was announced.

The directions only for tests 9, 12, 16, and 19 follow:

Test 9 Directions

You will hear many series of paired numbers. There will be twelve pairs of numbers in each series. Notice that this answer sheet provides for groups of twelve answers, and that each answer space within the group corresponds to a pair of numbers within the series. This is how and when you are to mark the answer spaces: When the first number of a pair is larger than the second number, mark the corresponding answer space plus (+). When the first number of a pair is the smaller and when the second number of this pair is an odd number, mark minus (—) in the appropriate answer space. The sample series below has been correctly answered, and the correct answers for Series W and X have also been indicated. You will notice that not every pair of numbers requires a response.

Test 12 Directions

By mixing digits and letters, six-item series have been prepared. A series may contain an equal number of digits and letters; it may contain more letters than digits, or it may contain more digits than letters. The digits may be any of the numbers from 1 to 9. The letters will be taken from the group A, B, C, D, E. There is an answer line for each series. In each answer line there is a space reserved for each letter of the series.

This is how and when you are to mark the answer spaces. If a digit is larger than the one preceding it in the series, mark the answer space for this digit plus (+). If a digit is smaller than the one preceding it in the series, mark the answer space minus (—). If a letter is one occurring later in the alphabet than the letter just preceding in the series, mark the answer space for this letter plus (+). If a letter occurs earlier in the alphabet than the letter preceding it in the series, mark the answer space minus (—). You will never mark the first answer space since that letter or digit will not be preceded in the series.

Test 16 Directions

Lists of digits will be read. You will respond to some of the dig-

its by marking a plus (+). You will respond to other digits by marking a minus (—). Some of the digits will require no response. This answer sheet has been prepared for recording your responses. There is an answer space for each digit.

This is how and when you are to mark the answer spaces below: Respond to all even digits by marking a plus (+) for the first even digit, by marking a minus (—) for the second even digit, by marking a plus (+) for the third, and so on, alternating plus and minus marks for the even digits. Mark odd digits plus (+) whenever the odd digit immediately follows an even digit. Do not respond to odd digits not following an even digit. Be sure that you place the response to a digit in the answer space reserved for it. Study the examples.

Test 19 Directions

In the answer space below, you are to write the letters A, B, C, D, E. Each time you will write these letters in a different order. You will be instructed in two stages. First you will be told to interchange the position of two letters; this will destroy the A, B, C, D, E, sequence. Then you will be told to place one letter between some two letters. At the signal you are to write the newly formed series in the answer space. Hold your pencils in the air as the instructions are given. You will have only time enough to write the series. Immediately you will be asked to raise your pencils, and you will hear new instructions for writing the letters another time.

Here is an example (remember that each time you start with the letters in this order— A, B, C, D, E): "Interchange A-B, now place E between A-C. Write the new series, mark!" At the word "Mark!", you will write down B, A, E, C, D. You will be given a few practice series. Make no sound or movement that can distract your neighbor.

It is important that the subjects' scores do not depend upon the readiness with which they learn the complex directions. This possibility would defeat the purpose of the tests inasmuch as they were constructed to emphasize following directions—not learning them. It was necessary, therefore, that every subject understand the task thoroughly before actual testing was begun. When the errors within each test were tabulated series by series or list by list, no tendency for the number of correct responses to increase was observable. On this basis it was assumed that the scores depend very slightly, if at all, upon the ability to learn the directions during the actual testing.

Since the primary purpose of the investigation was to ascertain whether there exists an "attention" factor, it was important to in-

clude in the total battery tests of other well-known factors which might possibly be confused with an attention factor. Factors which were suspected as possibly contributing to scores on the prepared tests were those of number, rote memory, visual-space, and perception. Due particularly to the investigations of Thurstone, convenient tests of these abilities were available. In the present study, tests 5 and 18 are tests of the number factor (7); tests 3 and 14 are tests of the rote memory factor (7); and 1 and 13 are tests of the visual-space factor (7). Cancellation tests have been found to contribute to the factor of perception (1, 8, 11, 12). Two cancellation tests, 6 and 10, and two tests that have been employed as tests of the perceptual factor, 2 and 7 (8), were included in the test battery. The problem thus became one of whether a factor were present in the newly-devised attention tests which was distinguishable from the other already well-established factors of number, rote memory, visual-space, and perception. If such an additional factor were found, it seemed unlikely that it could be identical with any of the other factors already well-established; so it seemed unnecessary to include tests of any additional factors.

In addition to the newly devised tests intended to be tests of attention, certain older tests were also included because of the possibility that they might involve an attention factor. These tests comprised the Otis Directions, number 8, Philip's Alphabet, number 15, both of which were found to be tests of attention in a study by a previous investigator (11, 12), and also the Seashore Tonal Memory Test, number 4 (5). Finally, for what now appears to be no good reason, there was included in the battery a test of syllogisms, number 20.

For the convenience of the reader, the identification of the tests included is summarized below:

Variable Number	*Test*	*Variable Number*	*Test*
1	Cards	11	Specially prepared
2	Identical Forms	12	Specially prepared
3	Initials Recall	13	Figures
4	Tonal Memory	14	Word-Number
5	Addition	15	Philip's Alphabet
6	Scattered *X*'s	16	Specially prepared
7	Backward Writing	17	Specially prepared
8	Otis Directions	18	Multiplication
9	Specially prepared	19	Specially prepared
10	Cancelling	20	Syllogisms

The tests were administered to four beginning classes at the Air Corps Technical School at Chanute Field, Rantoul, Illinois. The men,

TABLE I
INTERCORRELATIONS OF TESTS

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1																				
2	.263																			
3	.166	.180																		
4	.022	.247	.037																	
5	.097	.109	.167	.216																
6	.198	.205	.074	.003	.157															
7	.175	.266	.164	.226	.284	.237														
8	.208	.309	.108	.213	.300	.238	.232													
9	.141	.332	.155	.231	.340	.157	.164	.354												
10	.210	.414	.221	.018	.180	.417	.355	.188	.224											
11	.211	.212	.112	.201	.406	.220	.144	.343	.581	.210										
12	.198	.198	.138	.315	.230	.119	.159	.301	.340	.219	.377									
13	.609	.183	.055	.107	.013	.104	.241	.244	.097	.166	.170	.213								
14	.195	.129	.408	–.041	.233	.070	.191	.062	.117	.172	.125	.217	.080							
15	.165	.027	.029	.335	.342	–.024	.194	.320	.335	.268	.300	.268	.121	.181						
16	.097	.143	.134	.146	.178	.053	.122	.227	.307	.124	.418	.191	.040	.072	.220					
17	.159	.154	–.012	.113	.253	.179	.161	.356	.372	.205	.441	.381	.137	.173	.191	.376				
18	.099	.245	.192	.058	.585	.157	.246	.229	.501	.323	.316	.302	.126	.206	.213	.269	.084			
19	.251	.228	.037	.082	.200	.207	.100	.401	.256	.281	.357	.244	.184	.047	.271	.308	.334	.120		
20	.055	.008	.037	.213	.133	.053	.038	.259	.147	.110	.289	.213	.043	.100	.204	.025	.160	–.012	.204	

ranging in age from 18 to 30 years of age, had completed four years of high school or the equivalent and were selected on the basis of high standing in the Army Alpha for instruction in radio. Complete test data were obtained for 175 men.

All tests were administered in the testing room of the Trade Test Department at Chanute Field. The size of the classes ranged from 40 to 60 men. The tests were conducted in 1940 by the writer; the first class was tested Thursday morning and afternoon as well as the following Saturday morning in the second week of February. The second, third, and fourth classes were tested on the same schedule during the months of March, April, and May, respectively. The tests were administered to each class in the same order, and, so far as known, under identical conditions. Each of the tests is identified by a number which indicates its order of presentation. Hourly rest periods were given, and the men were allowed to talk and smoke between tests. The testing room was well ventilated and well protected from outside disturbances. The students were courteous and seemed to be highly cooperative.

Due to the nature of the tests, it seemed likely that some subjects might resort to guessing. In scoring the tests, therefore, a correction for chance successes was made by simply diminishing the number of correct responses by a certain fraction of the number of failures. The correction fraction was based upon the number of responses possible for an item.

The intercorrelations of the tests, based upon 175 cases, were calculated by the Pearson product-moment method and are found in Table 1. The correlations tend to be small and correlations between tests which ordinarily correlate highly were only relatively high. For example, the correlation between the two number tests is .585, and the highest of all the correlations, which is that between the two tests of spatial ability, is only .609. Perhaps no completely satisfactory explanation for the tendency of the correlations to be small may be offered; but part of this effect may plausibly be attributed to the homogeneity of the group. The men were selected for instruction in radio on the basis of high standing on the Army Alpha, and it is not unlikely that the Army Alpha measures to varying degrees the same things that are measured by many of the tests in the battery.

The table of intercorrelations was subjected to a centroid analysis and although nine centroid factors were extracted, Tucker's criterion as employed by Thurstone (7) was the only evidence indicating that more than seven factors should have been extracted. The first seven centroid factors are shown in Table 2.

TABLE 2

THE FIRST SEVEN CENTROID FACTORS AND h^2

	1	2	3	4	5	6	7	h^2
1	.448	-.363	.422	-.189	.159	.170	.050	.603
2	.463	-.233	-.053	.148	.117	-.184	.310	.437
3	.305	-.304	-.240	-.146	-.290	.112	.052	.364
4	.334	.240	-.130	-.303	.225	-.245	.267	.461
5	.544	.100	-.307	.049	.118	.107	-.347	.547
6	.352	-.223	.163	.315	-.084	-.204	-.112	.362
7	.440	-.240	-.156	-.057	.091	-.177	-.029	.319
8	.575	.164	.147	.038	.130	-.170	-.064	.431
9	.622	.194	-.157	.190	.098	.151	.165	.546
10	.513	-.342	-.041	.186	-.150	-.347	-.096	.569
11	.653	-.323	.062	.145	-.084	.197	.086	.608
12	.543	.134	.047	-.149	-.074	.065	.074	.351
13	.385	-.303	.430	-.252	.324	.139	.060	.617
14	.342	-.247	-.145	-.236	-.335	.198	-.141	.426
15	.467	.266	-.170	-.243	.120	-.130	-.181	.437
16	.420	.211	-.059	.161	-.095	.144	.177	.311
17	.506	.286	.189	.133	-.192	.104	.063	.444
18	.526	-.081	-.416	.234	.211	.302	-.227	.700
19	.490	.164	.234	.162	-.055	-.132	-.033	.369
20	.279	.254	.147	-.182	-.136	-.181	-.126	.269

TABLE 3

ROTATED FACTOR MATRIX AND h^2

	I	II	III	IV	V	VI	VII	h^2
1	.112	.074	.257	.168	.699	-.029	.063	.606
2	.428	-.026	.076	.182	.174	.024	.427	.436
3	.155	-.006	.531	.045	-.020	.211	.112	.366
4	-.006	.445	-.057	.146	-.002	-.004	.486	.459
5	.132	.193	-.015	.284	.044	.642	.031	.551
6	.516	-.061	.010	.236	.154	.022	-.110	.362
7	.372	.196	.144	.049	.149	.235	.208	.321
8	.239	.297	-.120	.449	.207	.117	.107	.429
9	.092	-.023	-.018	.535	.068	.331	.370	.547
10	.682	.097	.191	.185	.077	.192	.022	.568
11	.002	.038	.042	.727	.083	.211	.165	.610
12	.018	.233	.197	.445	.135	.122	.173	.354
13	.037	.147	.128	.089	.743	-.042	.124	.616
14	.062	.092	.562	.107	.050	.276	-.077	.424
15	.033	.499	-.021	.262	-.011	.319	.128	.437
16	.014	-.070	.057	.488	-.041	.138	.215	.313
17	.029	.042	.062	.660	.050	.029	.022	.446
18	.155	-.119	.003	.214	.120	.758	.165	.700
19	.245	.141	-.068	.517	.128	-.005	-.012	.368
20	.030	.392	.031	.307	-.024	-.049	-.098	.262

Perceptual ? Memory attention Space Number ?

TABLE 4
ORTHOGONAL TRANSFORMATION MATRIX

	I	II	III	IV	V	VI	VII
I	.358	.274	.225	.668	.302	.369	.272
II	–.498	.289	–.471	.517	–.427	–.037	.010
III	–.046	.017	–.112	.294	.594	–.632	–.390
IV	.427	–.703	–.447	.279	–.153	.080	–.106
V	–.082	.102	–.654	–.297	.516	.236	.383
VI	–.652	–.506	.259	.162	.288	.373	–.061
VII	–.095	–.276	.153	.107	–.057	–.516	.781

A preliminary attempt was made to employ a rotational method which involved an extension of the test vectors (9). This method was abandoned in favor of the more conventional graphic orthogonal method (6), which in the present study offered the advantage of greater precision in rotation.

The final solution, Table 3, was based on an orthogonal rotation of the first seven centroid factors. The transformation matrix is given in Table 4. These rotated factors were observed to be strikingly similar to those found by the method of extended vectors. The fact that two entirely different rotational procedures yielded the same meaningful factors is considered to be evidence of the validity of the rotation. That the same meaningful rotated factors resulted in each instance, despite the fact that the number of centroid factors involved was eight in the first method and seven in the second, may be considered as further evidence that seven centroid factors account for essentially all of the common factor variance.

In order that the exact nature of the rotated factors given in Table 3 may be more conveniently apprehended, each factor will be discussed in terms of the tests which have saturations greater than .400 with the particular factor in question.

Factor I (Perceptual)

Test	Name	Loading
2	Identical Forms	.428
6	Scattered X's	.516
10	Cancelling	.682

Factor I is interpreted as being a perceptual factor. Tests 6 and 10 are both cancellation tests and test 2 is one which has been found by Thurstone (8) to be a test of the factor which he calls perception.

Factor II

Test	Name	Loading
4	Tonal Memory	.445
15	Philip's Alphabet	.499
20	Syllogisms	.392

It would seem unwise to hazard any statement as to the exact nature of the ability or performance that is responsible for the persistent clustering of tests 4, 15, and 20. It may be said that these tests involve the high degree of sustained effort characteristic of Factor IV to a relatively slight degree. No other interpretative statements seem to be justified on the basis of the present findings.

Factor III (*Memory*)

Test	Name	Loading
3	Initials Recall	.531
14	Word-Number	.562

Factor III is obviously the rote memory factor. The loading of .531 for the Initials Recall test and .562 for the Word-Number test, although not particularly high, are higher than those found by Thurstone, .476 and .512, respectively, in his study entitled *The Perceptual Factor* (8).

Factor IV (*Attention*)

Test	Name	Loading
8	Otis Directions	.449
9	Specially prepared	.535
11	Specially prepared	.727
12	Specially prepared	.445
16	Specially prepared	.488
17	Specially prepared	.660
19	Specially prepared	.517

Factor IV is the factor of attention. It was to establish the existence and to define the nature of this factor that the present study was undertaken. The tests which are highly saturated with Factor IV, tests 8, 9, 11, 12, 16, 17, and 19, appear to depend upon the application of sustained, concentrated effort. The factor common to these tests cannot be explained in terms of the sensory mode of presentation. Tests 9, 11, 12, 16, 17, and 19 are presented by means of sound, but test 8 is Otis Directions and is presented as a printed sheet. On the other hand, tests 4 and 15, though also presented by means of sound, show no significant loadings with Factor IV. There is evidence, furthermore, that the factor does not depend upon content. Test 8, Otis Directions, involves printed words and other types of symbols; test 12 makes use of digits and letter combinations; tests 17 and 19 call for responses to letters of the alphabet alone, and tests 9, 11, and 16 call for responses to digits only.

Factor IV is similar to Woodrow's factor of attention (11, 12). Particular evidence for this similarity is the fact that the Otis Direc-

tions Test, number 8 in the present study, contributes to the factor of Attention in each of his studies, as well as to Factor IV. General evidence of the similarity is the fact that the writer's concept of the nature of Woodrow's factor determined to a large degree the type of tests constructed for the present investigation. That Factor IV is identical with Woodrow's factor, however, while a possibility, cannot be determined from the present investigation.

Factor V (*Space*)

Test	Name	Loading
1	Cards	.699
13	Figures	.743

Factor VI (*Number*)

Test	Name	Loading
5	Addition	.642
18	Multiplication	.758

Spatial ability, Factor V, is identified by tests 1 and 13. Numerical ability was measured by tests 5 and 18, which define Factor VI. The space and the number factors are so frequently encountered that they require no interpretation here. It is interesting to note, however, that the presence of numbers or numerical symbols in a test does not necessarily result in a high correlation of that test with the numerical factor. For example, test 16 involves responses to numbers only, while test 15 involves writing letters of the alphabet in a certain order; but in Table 3 it is seen that test 16 correlates .138 with the number factor, while test 15 correlates .319.

Factor VII

Test	Name	Loading
2	Identical forms	.427
4	Tonal Memory	.486

Since tests 2 and 4, the only tests showing significant loadings with Factor VII, do not seem to have anything in common, it is impossible to ascribe any definite meaning to the factor.

The present factor analysis was applied to this fundamental question: Will the variance of the tests which are included in the battery as possible measures of attention be accounted for in terms of factors known to be measured by certain tests in the battery or will some new factor identifiable as attention contribute largely to the variance?

From Table 3, the rotated factor matrix, it is apparent that the variances of certain of the tests included as measures of attention involve the factors which were known to be present in the battery (space, number, memory, and perceptual) to a negligible degree.

The relative importance of the various factors in contributing to the variances of the attention tests may be clearly indicated by the following type of equation **(6)**:

$$1 = \sigma_j^2 = (A_{j1})^2 + (A_{j2})^2 + (A_{j3})^2 + \cdots + (A_{jn})^2 + (A_{js})^2 + (A_{je})^2,$$

in which j refers to the particular test, A is the loading of the independent factors 1, 2, 3, etc., in this test, s is the specific factor, and e is the error factor of the test. The specific variances and the unreliability were not determined in the present investigation and will be combined as uniqueness, which may be defined by the equation **(6)**:

$$U^2 = 1 - h^2,$$

in which U^2 stands for uniqueness and h^2 is the communality of the test.

The analysis of the variance of each attention test into linear components is shown below with uniqueness substituted in the equations for the specific and error variances:

$$\begin{aligned}
\sigma_8^2 = 1 &= .058P + .008II + .014M + .202A \\
&\quad + .043S + .014N + .011VII + .571U^2; \\
\sigma_9^2 = 1 &= .008P + .001II + .000M + .286A \\
&\quad + .005S + .110N + .137VII + .453U^2; \\
\sigma_{11}^2 = 1 &= .000P + .001II + .002M + .529A \\
&\quad + .007S + .045N + .027VII + .390U^2; \\
\sigma_{12}^2 = 1 &= .000P + .054II + .039M + .198A \\
&\quad + .018S + .015N + .030VII + .646U^2; \\
\sigma_{16}^2 = 1 &= .000P + .005II + .003M + .238A \\
&\quad + .002S + .019N + .046VII + .687U^2; \\
\sigma_{17}^2 = 1 &= .001P + .002II + .004M + .436A \\
&\quad + .003S + .001N + .000VII + .554U^2; \\
\sigma_{19}^2 = 1 &= .060P + .020II + .005M + .267A \\
&\quad + .016S + .000N + .000VII + .632U^2,
\end{aligned}$$

in which σ_8^2, σ_9^2, σ_{11}^2, etc., are the squares of the standard deviations of tests 8, 9, 11, etc., and are equal to the total variance of the tests. P, II, M, A, S, N, VII represent the perceptual, II (unidentified), memory, attention, space, number, and VII (meaningless) factors, respectively. U^2 is the uniqueness variance.

From these equations, which show exactly what proportions of the total variance of each test are accounted for by the various factors, it is apparent that two tests, 11 and 17, are particularly good measures of the factor here designated as an attention factor. No less

than .529 of the variance of test 11 and .436 of the variance of test 17 depend upon the attention factor. The value of these tests as measures of attention is further indicated by the fact that other factors contribute negligibly to the variance. The greatest amount of variance contributed to tests 11 or 17 by any factor other than attention is .045, the contribution of the number factor to test 11. Although the degree to which the variance of tests 8, 9, 12, 16 and 19 depends upon the attention factor is considerably less than that for tests 11 and 17, none of these tests, with the exception of 9, draws upon factors other than attention to any appreciable degree. Test 9 is an unsatisfactory test of attention; almost one half of its common factor variance is concentrated in the number factor and in Factor VII.

Tests 11 and 17 are not only better measures of attention than any of the other tests, but they measure no other factor found in the study. It is of considerable interest to note that these tests are quite dissimilar in content. Test 11 comprises numerical material only, whereas test 17 includes only letters of the alphabet. Due to the length of the test series and the nature of the operations, there can be little question that the factor is one involving a continuous and sustained application of mental effort. Naming the ability, the existence of which is indicated by Factor IV, is admittedly an arbitrary matter. It is believed, however, that the nature of the tests in which it is manifested and the historical background for the present investigation make the term attention an appropriate one.

Conclusions

1. A single factor, which has been termed attention, accounted for the common factor variance of tests which were constructed to involve a high degree of sustained effort.

2. The factor is independent of content. The two best measures of the factor were different in content; test 11 involves digits alone and test 17 involves letters only.

3. There is no evidence that the factor depends upon the mode of presentation. While most of the tests contributing to the attention factor were presented in an auditory manner, by means of phonograph, test 8, which showed a significant loading with the factor, was presented in a visual manner. On the other hand, tests 4 and 15, which were also auditory, showed no significant correlations with this factor.

4. The fact that the two separate rotational procedures yielded essentially the same factors is evidence for the uniqueness and validity of the rotations.

5. Three Thurstone factors, memory, space, and number, were clearly defined in the orthogonally rotated factor matrix.

6. A perceptual factor heavily loaded with cancellation tests was also defined in the rotated factor matrix.

7. Two factors, II and VII, showed no important correlations with the tests and no attempt was made to give them a meaningful psychological interpretation.

References

1. Coombs. C. H. A factorial study of number ability. *Psychometrika*, 1941, **6**, 161-189.
2. Holzinger, K. J. Preliminary report on Spearman-Holzinger unitary trait study, No. 6. Chicago: Statist. Lab., Dept. Educ., Univ. Chicago, 1935.
3. Otis, A. S. Otis group intelligence scale. Yonkers, N. Y.: World Book Company.
4. Philip, B. R. The measurement of attention. *Studies in Psychology and Psychiatry*, 1928, II, No. 1.
5. Seashore, C. E., Lewis, D., and Laetveit, J. G. Seashore measures of musical talent, 1939 Revision. Camden, N. J.: Educ. Dept. R. C. A. Mfg. Co., Inc.
6. Thurstone, L. L. The vectors of mind. Chicago: Univ. of Chicago Press, 1935.
7. Thurstone, L. L. Primary mental abilities. *Psychom. Monogr.*, 1938, No. 1.
8. Thurstone, L. L. The perceptual factor. *Psychometrika*, 1938, 3, 1-17.
9. Thurstone, L. L. A new rotational method in factor analysis. *Psychometrika*, 1938, 3, 199-218.
10. Woodrow, H. The measurement of attention. *Psychological Monographs*, 1914, XVII, No. 76.
11. Woodrow, H. The common factors in fifty-two mental tests. *Psychometrika*, 1939, 4, 99-108.
12. Woodrow, H. Factors in improvement with practice. *J. Psychol.*, 1939, 7, 55-70.

PSYCHOMETRIKA—VOL. 8, NO. 1
MARCH, 1943

AN ANALYSIS OF RANDOM AND SYSTEMATIC CHANGES WITH PRACTICE*

EDWARD B. GREENE
UNIVERSITY OF MICHIGAN

Six motor tests and six nonverbal tests were administered four times to the same subjects. Subjective reports of the subjects are discussed, changes in mean scores and in variability and score correlations from trial to trial are surveyed, and factor analyses of results on the first and fourth trials are presented and compared. Implications of the findings with respect to correction for attenuation are pointed out.

The Problem

Many articles have been written concerning retest-correlations and split-half correlations which assume that a low coefficient signifies the presence of only chance variations, either in the individuals tested or in the methods of observing and recording their performances or in both. This assumption underlies the formulas for correction for attenuation and the Spearman-Brown split-half prediction formula. The correction for attenuation, urged by Spearman (**1, 2**) and many others, may in reality be eliminating or reducing important trends of a systematic sort. Although it is doubtless true that random errors and variations in performance will tend to lower self-correlations, the opposite, namely, that a low self-correlation indicates chance variations, is by no means true. Systematic errors of measurement and systematic changes in performance will also lower self-correlations. In the repetition of a test, it is almost inevitable that systematic changes in performance will be present. Likewise the first and last halves of many tests call out different ability patterns. Hence it is important in a careful analysis to know what sorts of changes are affecting the results. There are two approaches to this problem, one descriptive and the other statistical. In the analyses reported here both approaches were used because they supplement each other. The

* Thanks are due to Mr. Searles and his staff at the Henry Ford Trade School, who gave both suggestions and aid in completing this work; to Dr. L. L. Thurstone, who gave suggestions for certain aspects of the analysis; and to L. R. Tucker and Robert Blakey, who made the final factorial analyses and suggested various conclusions.

two prerequisites for a careful study of practice effects are generally agreed to be unambiguous measuring instruments, and a constantly motivated group which remains intact. These prerequisites were secured in part by administering four forms of the *Michigan Non-Verbal Series* (Greene, 1931), one a day, to 394 boys in training at a large mechanical trade school. The boys ranged in age from 14 years and 2 months to 15 years and 8 months, mean 14 years and 9 months. Most of them had finished the eighth grade in the Detroit area and they were in general selected from the upper half of the scholarship ratings.

The *Michigan Non-Verbal Series* has four forms which have been carefully equated for difficulty at two or three levels of complexity. The series includes twelve sub-tests which were always given in the following order:

1. *Aiming*: with pencil, left hand, at 4mm. circles for one minute.
2. *Aiming*: with pencil, right hand, at 4mm. circles for one minute.
3. *Aiming*: with pencil, left hand, at 1½ mm. circles for one minute.
4. *Aiming*: with pencil, right hand, at 1½mm. circles for one minute.

(Each aiming test was preceded by a 20-second warming-up work period. A standard No. 2 pencil of standard sharpness was used. All subjects were required to proceed from left to right across a row of five circles, then proceed from left to right on the next row below, until five rows were completed.)

5. *Tapping*: with pencil, left hand, 10 seconds, best of three trials.
6. *Tapping*: with pencil, right hand, 10 seconds, best of three trials.
7. *Feature Discrimination, Easy*: to circle with a pencil the unique feature in a row of *three* small printed faces, otherwise identical, 2-minute period.
8. *Feature Discrimination, Medium*: to circle one unique feature in one of *five* small faces printed in a row, 3-minute period.
9. *Feature Discrimination, Hard*: to circle with a pencil the unique feature in one of *seven* small faces printed in a row, 5-minute period.
10. *Maze Solution, Easy*: to draw a pencil line through rows of small square printed mazes, each having five alleys per side, without crossing any printed lines, 2-minute period.

11. *Maze Solution, Medium*: to draw a pencil line through square printed mazes, with 20 alleys per side, without crossing any printed lines, 3-minute period.
12. *Maze Solution, Hard*: to draw a pencil line through rectangular mazes ten times as large in area as those in No. 10, without crossing any printed lines, 5-minute period.

Four trials were given using the standardized equivalent forms, A, B, C, and D of these tests. Only the scores of boys who completed every trial of each test were included. The test scores were the number of correct responses in a period of time which was thought to be long enough to give a good sample of performance but not long enough to cause fatigue.

The tests were given in small class rooms during regular school periods in both morning and afternoon. Each group was tested on four consecutive days at the same time of the day. The motivation and effort seemed high and fairly constant. The feature discrimination tests were rated as less interesting than the other types of tests by nearly 90% of the entire group. The effects of interest on the scores have not been evaluated.

Descriptive Analysis of Processes

In order to give a clear picture of processes involved in success in the test situation, a summary of reports by subjects and observers is given. These reports were obtained from: 1) interviews with thirty subjects who were chosen to represent superior, average, and inferior performance on each type of test; 2) reports of six persons who observed trials of these thirty subjects, and others as well; and 3) reports of ten University of Michigan students who took the entire series of tests and wrote subjective analyses of their experiences after each trial. These reports indicate changes with practice toward the elimination of unnecessary movements and of acts usually associated with emotional behavior, such as smiling, scowling, coughing, etc. Many references are found to greater calmness in the later test situations, although none of the subjects were regarded as unusually emotional. An attempt was made to secure good motivation by telling the students that fair samples of their ability were desired and that each would be given a report of his test results at the conclusion of the series.

On the initial aiming tests the dominant processes seemed to be a flexible coordination of forearm and finger muscles and a rhythmic grouping of circles into horizontal sets and pauses. The more rapid and accurate students used very small finger and wrist move-

ments which gradually and directly approached the circles, with a well-marked rhythm of four or five dots, then a pause. The slower and more inaccurate students used large forearm and shoulder movements instead of, or alternating with, the finger and wrist movements. They exhibited irregular rhythms with two or three dots between pauses. They made the pencil point approach the circle with a ballistic or hammering movement which was sometimes checked and redirected before reaching the paper. They made heavier marks on the paper and more slips after placing the pencil than the faster students. The final trials showed a marked tendency toward smaller, more direct, less forceful movements, more regular rhythms, larger groups of circles, and shorter pauses between groups. The right-handed persons showed the heavy, ballistic, irregular forearm movements with the left hand, and reported greater effort and fatigue in left-hand aiming. The smaller circles required considerably more precision of movement than the larger ones.

On the initial tapping tests ballistic wrist and forearm movement seemed universal. The subjects were instructed to let the wrist lie comfortably on the table, but this was difficult for many students, especially with the non-dominant hand, and for slow subjects with either hand. The non-dominant hands used longer strokes, and hit the paper harder and more irregularly than faster hands. The shortest strokes were raised approximately one-quarter inch above the paper, and the longest, two inches. With practice, both hands tended to use shorter, lighter, and more regular strokes. No regular rhythms could be observed with either hand, and no spatial patterns were found. Students were asked not to try to tap a pattern. Most of them succeeded in letting the taps fall in a random fashion over an area of approximately one and three quarters inches in diameter.

The main process involved in the initial trial of the easy feature discrimination was a visual comparison of pattern. This was not dependent upon fine visual acuity. Each pattern was printed in large black lines and differed very noticeably in only one of the following ways: position, curvature, and length of line, or size and position of the dots representing eyes. The most successful students first compared all three faces with respect to one feature, such as hair. If the hair in all three was alike, some other feature, such as mouth, was selected and compared. This selection, comparison, and rejection was continued until the unique feature was found. The student then circled it with a pencil and proceeded to the next problem. Success in this test depended upon rapid and systematic isolation of a feature for comparison and upon rapid comparison of the three faces. Poor

scorers tried to compare whole faces, or repeated the comparison of features after they had been found to be uniform. The commonest error was the circling of one of two similar features. This error was due, according to subjects' reports, to a failure to compare all three faces. Some of the poorer scores were made by persons who isolated an attribute, such as length, and made a comparison of all features on this basis, overlooking for the time being any other differences such as curvature, position, or size. Changes during practice were generally in the direction of more rapid selection and isolation. more methodical comparison, and of the inclusion of more than one attribute of a feature in a comparative act. The comparison order changed from a somewhat random one, to a left to right sequence for nearly all subjects. The act of circling the unique feature with a pencil became more rapid with practice and a little less precise.

The medium feature discrimination tests included the same processes as the easy test just described. The muscular acts seemed nearly the same for all levels. The main differences between levels of difficulty are all reported to be in comparison processes. The comparison of a row of five faces demanded a more extended span of comprehension and a more methodical comparison than that of three faces. Avoidance of the repetition of comparisons became much more difficult and important. A slight distraction or emotional blocking would often make it necessary for the student to begin a problem all over again. Practice effects were also noted in a calmer and more determined approach to the task, a greater familiarity with the possible patterns in the test.

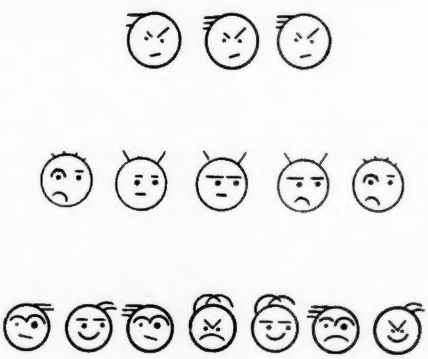

Figure 1
Samples from the Feature Discrimination Test

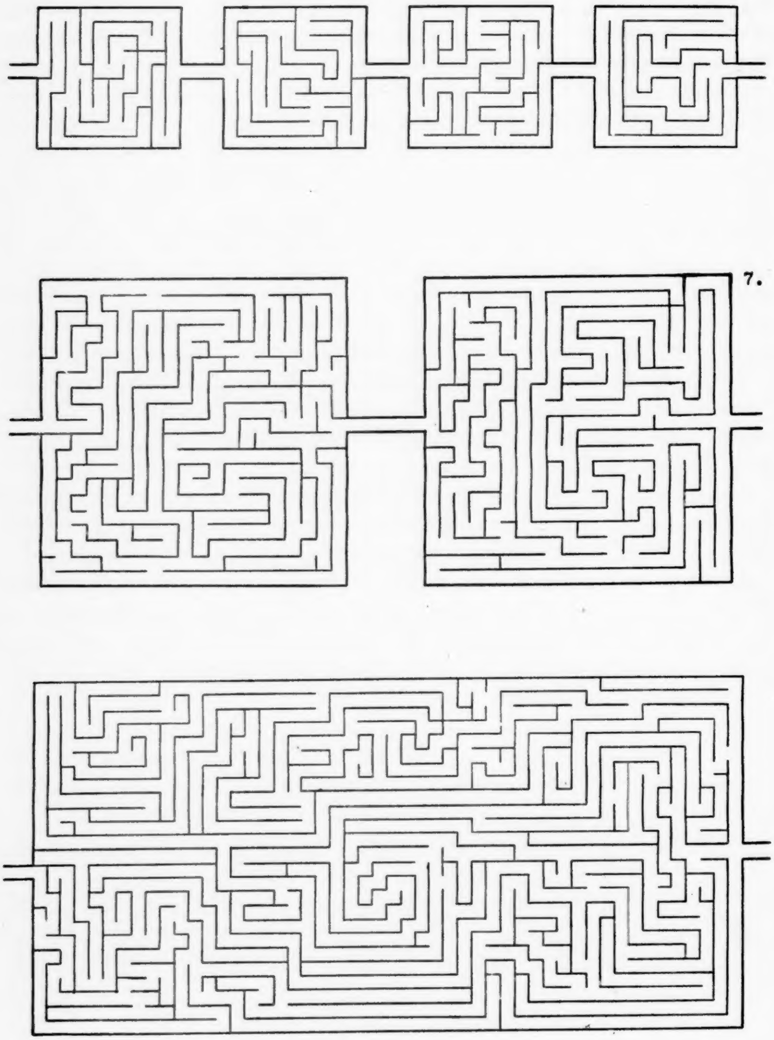

FIGURE 2
Samples from the Maze Test

The hard feature discrimination tests necessitated a comparison of a row of seven faces in which each feature always had three variations without showing a unique one. (Figure 1). The task then was to find a fourth variation which was different from any of the rest. Changes in skill seemed much like those noted above; the emotional elements and the concentration needed on each problem were much greater. A few students complained of fatigue and eye strain during the five-minute periods used.

The initial trial of the easy maze test seemed to be dominated by processes of simple visual comparison and of drawing lines. The comparisons of two pathways to see which led through to the next junction or to the end of the maze were not difficult for the more able subjects, since the blinds averaged 4 cm. in length. (Figure 2). Their test scores depended largely upon speed of drawing a line through alleys where the true pathway was perceived at a glance. Those making low scores may be divided roughly into two classes, the cautious, and those poor in comparison. The cautious persons generally made few errors and showed a more rapid improvement with practice than the others. Among those poor in comparison there was a considerable change with practice toward a longer comprehension span, fewer repetitions of entries into the same blind, more alertness in detecting openings and fewer emotional disturbances when no true path could be quickly found.

The medium maze test contained much more difficult comparisons, since the blinds averaged 8 cm. in length. Most of the students could not take the longest blinds in at a glance, but had to draw a pencil line some distance before discovering the dead end. Back tracing and repeated entries into the same blind were common at first, but were gradually reduced with practice.

The hard maze test differed from the medium in having greater length of blinds, averaging 12 cm., and greater complexity of pattern. With the larger area it was possible to make the true path follow spirals and to go long distances in the direction opposite from the goal, both of which were sources of error.

With practice, a skill was developed in avoiding the longer blinds, by means of visual explorations, and avoiding the repeated entries into blinds. Speed of drawing seemed to have little effect upon the score and it sometimes would lead to wrong decisions. The students were told to try to avoid blinds in order to get through the maze, but there was no penalty attached to errors other than the necessity of withdrawing from a blind. From these descriptive analyses, it seems fair to conclude that changes in performance from one trial to the

TABLE 1

Four Trials of the Michigan Non-Verbal Series, Amounts Correct; Raw Score Means, Standard Deviations, and Coefficients of Variation

Form A—1st Trial

Test	Mean		S.D.	C.V.
Aim Left, Large Circles	49.57		12.21	24.63
Aim Right, Large Circles	52.54		20.25	38.54
Aim Left, Small Circles	26.94		10.02	37.53
Aim Right, Small Circles	45.38		18.35	42.52
Tap Left	46.02		14.04	30.45
Tap Right	57.18		18.44	31.81
Feature, Easy	14.23		6.44	45.47
Feature, Medium	6.66		2.73	40.99
Feature, Hard	5.36		3.61	67.72
Maze, Easy	14.15		4.54	31.34
Maze, Medium	18.81		7.14	37.96
Maze, Hard	9.56		9.92	107.05

Form B—2nd Trial

Test	Mean	% Improvement	S.D.	C.V.
Aim Left, Large Circles	51.81	4	13.8	26.63
Aim Right, Large Circles	71.26	36	16.55	23.27
Aim Left, Small Circles	31.10	15	8.68	28.05
Aim Right, Small Circles	48.35	6	12.56	26.10
Tap Left	51.21	11	8.46	20.55
Tap Right	59.94	5	4.32	7.33
Feature, Easy	19.58	38	5.88	29.72
Feature, Medium	9.40	42	3.77	40.11
Feature, Hard	6.35	18	2.11	48.57
Maze, Easy	17.98	26	5.56	31.06
Maze, Medium	18.31	–3	7.82	42.71
Maze, Hard	14.62	53	7.74	53.08

Form C—3rd Trial

Test	Mean	% Improvement	S.D.	C.V.
Aim Left, Large Circles	60.07	21	16.10	26.81
Aim Right, Large Circles	72.46	38	18.0	24.84
Aim Left, Small Circles	32.86	22	10.98	33.20
Aim Right, Small Circles	50.25	11	12.12	24.01
Tap Left	53.91	17	9.06	16.72
Tap Right	63.02	11	8.28	13.26
Feature, Easy	21.23	50	6.02	28.33
Feature, Medium	12.58	99	4.68	37.25
Feature, Hard	9.47	76	3.90	41.17
Maze, Easy	17.92	26	5.84	32.41
Maze, Medium	24.05	28	8.79	36.55
Maze, Hard	20.28	112	10.62	52.28

Form D—4th Trial

Test	Mean	% Improvement	S.D.	C.V.
Aim Left, Large Circles	60.42	22	15.25	25.24
Aim Right, Large Circles	77.35	46	19.72	25.48
Aim Left, Small Circles	33.06	22	10.38	30.64
Aim Right, Small Circles	51.9	14	13.24	24.43
Tap Left	55.9	21	10.62	19.38
Tap Right	72.44	26	12.40	17.37
Feature, Easy	22.73	59	6.34	27.93
Feature, Medium	14.14	113	4.28	30.27
Feature, Hard	8.66	61	3.62	48.01
Maze, Easy	19.28	35	5.50	28.37
Maze, Medium	25.63	38	10.53	41.08
Maze, Hard	20.30	114	11.13	50.02

next were to a large degree qualitative. Even in the simplest motor test, tapping, there was a change in the use of muscles holding and moving the pencil which can hardly be described as random, since there was a definite pattern of shift from large to small muscle groups. Evidence of changes of a random sort were principally small variations in the expenditure of energy in a given time.

Statistical Results

In evaluating the random and systematic variables in successive trials of a test, no single measure is sufficient. It is necessary to show how changes in self correlations are related to changes in raw scores, and to change in factor loadings. Table 1 shows improvements in raw score means, and standard deviations. Table 2 shows the intercorrelations of various trials of the same test, and Table 3 shows a factorial analysis of the 1st and 4th trials.

Improvements in Achievements

An inspection of Table 1 shows an increase in both means and S.D.'s with practice for all tests. In general, simpler intellectual tasks show smaller relative increases. Thus, aiming and tapping tests have a median increase in raw scores of 22% in four trials; that of feature and maze tests is 60%. Furthermore, hard maze and feature tests show greater relative increases than do easy forms. From subjective analyses above it seems clear that improvement on these trials depends upon selection and formation of routine procedures. When a routine pattern is well practiced, as in tapping with the dominant hand, little improvement is shown.

It also appears that the S.D.'s increase proportionately less than the means, so that the coefficients of variation ($C.D. = \frac{S.D.\ 100}{\text{Mean}}$) become smaller with practice. This table illustrates a general rule that tests showing smaller relative increase also have the smallest initial C.V.'s and the smallest changes in C.V.

These results lead one to examine the zeros of the scales. Obviously the raw score zeros do not represent just no ability, hence the true zero will be below the zeros of these scales. Thurstone (3) has postulated a true zero as that point where the dispersion of a normal group is zero. Our data here represent age ranges that are too small to justify an application of his formula, but on the basis of previous work, (4), it seems that all of these tests tend to have similar C.V.'s at approximately 14.7, when normal samples of various ages are

TABLE 2

Michigan Non-Verbal Series, Amounts Correct; Product-Moment Correlations Between Trials of Equivalent Tests Uncorrected for Attenuation

Test	Trials	Trials 2	Trials 3	Trials 4	e^2 Begin	e^2 End
1	1	.531	.557	.457	.720*	
Aim Left	2		.676	.745		
1½ mm.	3			.793		.372
2	1	495	.374	.290	.757	
Aim Left	2		.552	.451		
4 mm.	3			.782		.390
3	1	.397	.541	.443	.844	
Aim Right	2		.587	.633		
1½ mm.	3			.772		.406
4	1	.479	.268	.231	.762	
Aim Right	2			.569	.531	
4 mm.	3			.581		.663
5	1	.369	.303	.243	.864	
Tap Left	2		.374	.256		
10 sec.	3			.438		.808
6	1	.421	.414	.336	.822	
Tap Right	2		.346	.272		
10 sec.	3			.312		.902
7	1	.532	.583	.421	.716	
Easy	2		.591	.692		
Feature	3			.614		.624
8	1	.426	.374	.375	.818	
Medium	2		.473	.323		
Feature	3			.361		.870
9	1	.233	.211	.257	.945	
Hard	2		.352	.313		
Feature	3			.428		.816
10	1	.726	.592	.592	.472	
Easy	2		.690	.647		
Maze	3			.881		.220
11	1	.504	.513	.602	.747	
Medium	2		.494	.501		
Maze	3			.534		.716
12	1	.224	.280	.226	.950	
Hard	2		.351	.182		
Maze	3			.534		.716
Median	r_{12}	.450	r_{34}	.560	.796	.688

* Calculated from equivalent-form retests one day apart.

measured from true zero points calculated by Thurstone's method. The differences seen here in C.V.'s are therefore probably due largely to differences in raw score zero points, which, in turn, are related to the complexity of the task.

Changes in Self Correlation

Table 2 shows that during practice correlations between equivalent forms of a test tended to be larger between successive trials. The median r_{12} is .450, while the median r_{34} is .560 for all tests. The increases in correlation were greatest in aiming, hard feature, and hard maze tests, and smallest in tapping and medium feature and medium maze tests. The biggest correlations between 1st and 4th trials were found in aiming, medium and easy feature tests, and easy maze tests. There was a slight tendency for tests which had lower initial self correlations to have the larger increases with practices. This was particularly true of hard feature and maze tests and aiming with the dominant hand at larger circles. It was not true of tapping with either hand. Tapping was the test which allowed the greatest number of short variations in timing, because of the short time interval used, 10 seconds.

Factorial Analyses

The correlation matrices of the 1st and 4th trials were analyzed by Thurstone's (5) center-of-gravity method, using lengthened vectors. No effort was made to impose orthogonality among the planes, but every dimension was well represented.

The structure which appears from the five factors shown in Table 3 is one of the clearest seen from such data. Four planes, A, B, C, and D are very well defined, and E is fairly well delineated. All may be considered to be experimentally uncorrelated in the group tested.

An inspection of Table 3 shows that the total variances which were due to common factors, h^2, were usually larger in the fourth than in the first trial. This means that group factors were more important after practice than before. We are not justified in claiming that the factors listed in both trials are the same. This claim could only be made if both trials had been included in a single factorial analysis. However, it seems probable from the subjective reports earlier in the article that the important changes in processes which accompanied practice on various tests can be described in terms of the factor loading changes found in Table 3. The two factorial analyses will be compared, and the factors tentatively named.

TABLE 3

Effects of Practice on Factorial Analyses (Greene, 1938, Michigan Non-Verbal Series)

Test	*Form A. 1st Trial*						*Form D. 4th Trial*					
	A	*B*	*C*	*D*	*E*	h^2	*A*	*B*	*C*	*D*	*E*	h^2
1. Aim Left, Small	.10	.73	.00	.07	.00	.55	–.05	.46	–.08	.79	–.01	.83
2. Aim Right, Small	.02	–.09	.35	.69	.01	.61	.25	–.34	.01	.51	.06	.44
3. Aim Left, Large	.08	.72	.54	.12	.00	.83	.04	.34	.07	.75	.03	.68
4. Aim Right, Large	.08	–.03	.30	.64	.06	.51	.04	–.04	.41	.58	–.03	.51
5. Tap, Left	.01	.14	.57	.03	.01	.35	.05	.66	.40	–.01	.07	.60
6. Tap, Right	–.06	–.05	.56	.09	.02	.33	.01	.03	.76	–.03	–.07	.58
7. Feature, Easy	.47	.12	.34	.14	.12	.38	.73	–.11	.18	.04	.23	.63
8. Feature, Medium	.46	.08	–.11	.12	.19	.38	.70	.01	–.01	–.02	.05	.49
9. Feature, Hard	.69	–.06	.16	.05	.02	.50	.72	.12	–.09	–.06	–.16	.55
10. Maze, Easy	.04	.29	.06	.00	.68	.55	.07	.37	.05	–.01	.75	.58
11. Maze, Medium	.02	.32	.37	–.10	.71	.76	.00	.07	.04	.02	.71	.66
12. Maze, Hard	.09	–.19	.11	–.02	.39	.21	–.03	–.08	–.12	–.05	.75	.58

A. Feature Discrimination (isolation disconnected visual patterns)

B. Ambidexterity or persistence

C. Tapping and easy aim. Both hands. (Speed of ballistic movement)

D. Aim. Both hands. (Precision of movement or rhythm)

E. Maze solution. (Trail finding)

On the first trial the left hands have a very high loading in factor B, and small loadings in D. In the 4th trial these loadings are nearly reversed. The right-hand performances in both trials of large and small circles are nearly zero in B and very high in D. Thus the left hands are seen to change toward the right-hand patterns. This change was also noted in the subjective analysis. Two of the aiming tests, Right small, and Left large, show a reduction with practice in factor C, which is closely identified with speed of tapping. Aiming, Left small, shows none of factor C in either trial, while aiming, Right large, shows moderate amounts of C in both trials. This finding corresponds to observation that ballistic movements were usually reduced with practice except in the case of aiming, Right large. There precision had been developed to a point where nearly automatic movements were effective.

Tapping tests with the left hand show small amounts of factor B and large amounts of C on the first trial. Tapping Right shows no B but a large C loading. On the 4th trial Tapping Left has a larger amount of B factor than formerly and a smaller amount of C, while Tapping Right has no B and more C than formerly. The tapping factor patterns of the two hands, therefore, are shown to be less alike

after practice than before, although tapping performance was more alike as practice continued. This analysis indicates that after practice the variations in tapping of the two hands were apparently related to factor B more than the variations in aiming.

In the 1st trial all the feature tests are weighted most heavily by factor A, but they also show small amounts of other factors, particularly C. In the 4th trial the Feature Discrimination tests are all heavily loaded with factor A and show no other significant loadings. A very small amount of factor C, which corresponds to speed of rough marking, is still found in the easiest test. With practice the easiest feature discrimination test becomes in its factorial patterns more like the hardest test. From the subjective analysis it appeared that practice eliminated unnecessary movements by establishing regular methods of comparison.

The maze tests give much the same factorial picture. In the 1st trial they are all heavily loaded with E but they also have small amounts of B and C. In the 4th trial they are all heavily loaded with factor E and show no other significant loadings.

The identification of the statistical factors with behavior patterns is justified to the degree that competent judges agree upon the analysis of processes. Since in this case the tests were fairly well controlled and limited to particular skills, the subjective identifications are fairly clear, with one exception, factor B. Factor B may be allied to ambidexterity, since it appears principally in left-hand Aiming and Tapping, and easy Maze tests, performed by persons who were right-handed in 90% of the cases. Factor B may also be an indication of persistence in a speed trial. Factor A is a form of non-verbal induction using visual patterns. It is found only in the Feature Discrimination Tests. Factor C corresponds most clearly to speed in a series of small ballistic movements, as in tapping. Factor D is identified with speed of precision of small hand and eye movements, coordinated as in aiming. Factor E is found only in the pencil mazes. It is therefore identified with isolation of a visual path in a connected visual pattern.

Discussion

Evidence of Systematic Variations

Systematic changes in behavior may be described as fairly persistent changes in patterns of response to standard stimuli. Such changes are often attributed to different combinations of various mechanisms, but they may be changes in energy only. Random changes are usually described as changes which persist only for very

short periods, and hence do not affect the person's usual mode of response. Systematic variations would be indicated by changes in group norms, while random changes might be present if there were no changes in group norms.

The subjective reports above indicated that some of the tests showed more evidence of systematic changes than others. The two tests of aiming with the right hand and tapping with each hand were reported to have undergone only small changes in procedure with practice, while in all the other tests large changes occurred. The short tapping and aiming tests were reported to fluctuate more than the rest, because of difficulties in starting and stopping promptly and because of variations in energy or willingness to work.

In the statistical data a systematic change is indicated when, in several trials of a test, individuals' scores show a series of changes in a given direction. Hence systematic changes would be indicated by a change in the mean scores in Table 1, a change in correlations between trials in Table 2, or a change in factor loadings in Table 3.

Table 1 shows that all tests changed in mean scores, but only small variations were found in Aiming, Right small. Each of the other tests showed at least 20 per cent improvement over the first score.

Table 2 shows that only five tests showed small changes in the size of correlation coefficients between adjacent trials. These tests include Aiming, Right large, Tapping, Left and Right, Medium Feature Discrimination, and Medium Pencil Maze. Hence systematic changes were probably present in the other test situations.

Table 3 shows only one test, Aiming, Right large, which has practically the same factorial loadings in the 1st and 4th trials. One must conclude that only this test shows no marked evidence of systematic changes. There may have been systematic variations present in performance, but none appear in this analysis of scores. All the other tests have rather marked changes which probably indicate systematic variations in individual patterns of response.

The Most Probable Limits of Random Variation

Thurstone (5), Garrett (6), and Guilford, (7) have used e^2 in the formula

$$e^2 = 1 - r^2_{11}$$

to indicate the proportion of the total variance which is called random, in a special situation where it can be assumed that no systematic changes have taken place in two trials 1 and I of a test. Since this assumption has been shown to be untenable in most of the tests under

consideration, what value may be assigned to random variance? No direct answer appears to this question. If one accepts the hypothesis that random factors will be at least as great at any point during the practice as they will be at the end of a long practice series, then the self correlation at the end of practice will indicate a minimum e^2. This is a reasonable assumption in almost any series, because individual consistency usually improved with practice. For the same reason the maximum true e^2 for any point in a practice series will probably not be greater than that actually found somewhere near the beginning of the series. The maximum e^2 will usually be less than that indicated, for systematic changes, which eventually may become fairly permanent, can never be discerned from two adjacent trials alone. The e^2's shown in Table 2 are therefore probably higher than they should be, with the exception of Test 4, which seems to have reached a fairly stable pattern for the persons tested.

The Significance of the Quantitative Data

The quantitative data present several significant findings since they represent behavior patterns of a fairly large group of adolescent boys. First, the relative mean and standard deviation increases are large enough to make it necessary to compensate for practice effects in practical comparisons of individuals at various stages of practice. Second, the stability of factor loadings of one aiming test and the lack of stability on the other three aiming tests and on the tapping tests are of interest to those who seek measures of lateral dominance. Third, the factorial analyses are of interest to those who are approaching the understanding of mental organization from this angle. The results show quantitative indicators of changes in mental organization which take place during practice, and also indicate that most of the changes which were reported subjectively can be well represented as the resultants of five factors. The clarity with which these factors are isolated statistically gives promise that, from tests which are somewhat better controlled than these, very pure measures of abilities may some time be made which will be relatively free from factorial changes during practice. The nature of the five factors is also of interest. Evidence is presented here that ballistic and precision movements are unrelated and that isolation of a trail from a maze and isolation of a feature in one of several faces are two unrelated forms of selective activity.

Summary

Six motor tests and six non-verbal tests of observation and comparison were applied four times in one week to 394 adolescent boys

enrolled in a trade school. Observations and subjective reports of changes with practice were compared with quantitative results. The more complex tests showed more improvement with practice, more reduction of C.V.'s, and less predictability than did the less complex tests. Subjective reports indicated more varieties of approach to the complex than to the easy tasks and more elimination of useless acts among the complex tests by the methodical habit forming.

The first and last trials showed unrelated factors which are tentatively identified with:

a) Isolation of unique visual pattern from series of disconnected patterns. (Feature discrimination)
b) Ambidexterity. (Left-hand aiming and tapping, and easy maze solution)
c) Speed of ballistic movements. (Tapping)
d) Speed of precision movements. (Aiming)
e) Isolation of visual path from larger pattern. (Maze)

All except one test showed marked shifts in the size of factor loadings with practice. The shift was usually toward a larger loading on one factor and smaller loadings in all the rest. The presence of one test (Aiming, Right hand, Large circles) which maintained one factor pattern throughout the practice series, points to the possibility that more tests with constant factor patterns may be constructed.

The results lead to two suggestions. (a) Corrections for attenuation should be limited to situations where no systematic changes occur between trials as shown by factor loadings and subjective reports. (b) Since reliability coefficients may be indicative of both systematic and random changes, they should not be used routinely in corrections for attenuation. The probable limits of random variation can be described from a long practice series.

PSYCHOMETRIKA—VOL. 8, NO. 1
MARCH, 1943

ON THE INTERPRETATION OF COMMON FACTORS: A CRITICISM AND A STATEMENT

H. A. REYBURN AND J. G. TAYLOR
UNIVERSITY OF CAPE TOWN

The concept of simple structure is criticized for lack of objectivity and for failure to produce invariance of (*a*) factor loadings under change of battery, and (*b*) individual scores on primary traits under change of tests. It is also criticized on the grounds that simple structure yields at best the factors which were put in and that, by suitable manipulation of tests, any set of factors may be represented in a battery by tests which will yield a simple structure. A procedure for rotation is developed which locates the first rotated axis to pass through a cluster of tests which, by hypothesis, contain a common factor and to project all tests into a hyperplane orthogonal to this factor. The second factor is then located to pass through the projections of a second cluster of tests defining the second factor, and so on, until all hypothecated factors have been located. Any residual interrelations may then be rotated graphically to the most plausible arrangement.

The purpose of this paper is to discuss certain fundamental principles of factorial analysis in its application to problems of human behavior. The general question of the relation of factorial methods to other forms of psychological explanation does not concern us at the moment; we assume that factorial analysis has a legitimate field in psychology, although we do not assume that this field is coextensive with personality or behavior as a whole. Thus, for example, it may be asserted that every person, and ultimately perhaps every human act, is unique and cannot finally be accounted for by general laws which presuppose uniformities of nature and behavior. We do not deny this, but we would point out that uniformities are found and that the uniqueness and difference of personalities do not exclude observed agreements. The problems which methods of factorial analysis are employed to solve begin from observed agreements, and only by observation and experiment can the extent of these agreements be discovered. General theory will enter in and only in the light of it can human behavior be made intelligible. But it must not be introduced too early. We must first get the facts, and let the theory follow.

The starting point of factorial analysis is certain observed agreements in human behavior, usually expressed in the form of correlation coefficients. These coefficients, if significant, require interpretation.

The principles in terms of which this interpretation is given are usually called common factors, but no special implication should be read into this term. It is not implied at the outset that common factors are causes, or in any intelligible sense the immediate sources of the observed agreements. Primarily they are mathematical functions in which the agreements are reduced to their simplest terms, and *per se* they are nothing more than that.

Several techniques have been devised for the discovery of such factors; in what follows we presuppose the use of Thurstone's centroid method as the simplest and most convenient.

Conception of Objectivity

Not all common factors are of equal value for psychological purposes. Behind agreements in behavior there are sources of agreement, and psychological investigation must try to reach these sources. That is to say, if possible, explanation should be causal. But it is not always possible to attain this ideal, and something inferior may have to be accepted. If common factors are not causal, they must at least be objective. This does not mean that they must be ultimate or physical or concrete, but it does require a certain form and degree of invariance. The demand may be expressed in two propositions:

(a) The common factors which enter into a variable, such as a test of intelligence or temperament, are not altered when the battery in which the test is included is changed by the addition or exclusion of other tests.

(b) The quantity of any common factor manifested in the behavior of an individual in a given situation remains the same whatever tests be employed to measure it.

It is not necessary to elaborate these principles. They are not new and they have been stated quite clearly elsewhere. We submit, however, that their importance has not always been recognized, and that the existing techniques do not always take them fully into account.

If a factor is not objective in the sense defined above it is almost entirely valueless for psychological purposes. It is useless for prediction, because *ex hypothesi* the measure of it in a test situation is no clue to the measure of it in any other situation about which prediction has to be made. Factors lacking objectivity are inferior to the original tests from which they are derived, for these tests, whatever their deficiencies, are at least objective. This consideration, of course, has been given some weight, for it is ultimately on such grounds that Thurstone has denied psychological significance to the unrotated cen-

troid matrix. Nevertheless, current theory does not go far enough, and sometimes the methods adopted to rotate a centroid matrix to a significant position are in conflict with the demand for objectivity from which the whole process of rotation arises. We refer here chiefly to the conception of simple structure which has been widely used to guide the rotating axes to their final positions.

Simple Structure

Two reasons may be given for preferring a simple structure solution of the problem of rotation. Firstly, it is in harmony with the principle of parsimony: on the average, each variable is explained by a smaller number of general principles than can be obtained by any other solution. And secondly, a simple structure is unlikely to arise by mere chance. We may begin with the second argument.

It is true that simple structure is not likely to appear by chance in a battery of tests, whether of intelligence or personality or any other aspect of behavior. If it can be extracted by analysis it must have been inserted into the tests at the time of their construction. This may arise with or without conscious choice on the part of the framer of the tests. If an investigator, either deliberately or implicitly, has a number of factors in mind, and in assembling a battery of tests to measure them selects tests which will enable the factors to be discriminated as clearly as possible, it is not unlikely that simple structure will be introduced into the battery. One qualification, however, must be made. If the subsequent analyses are made by methods which involve the orthogonality of the axes, the original factors guiding the choice of the tests must be mutually independent in the sense that they are uncorrelated in the population measured.

It may be argued, then, that if a centroid matrix can be rotated orthogonally to show simple structure, the result obtained yields the same independent factors as those which guided the experimenter in his construction of the battery. What has been taken out is what has been put in, and surely nothing more can reasonably be demanded. This conclusion, however, does not follow. Independence, in the sense defined, although commonly an ingredient of objectivity, is not identical with it and does not guarantee it. Perhaps the point may be put in this way. In an ordinary centroid analysis axes may be rotated orthogonally into an infinite number of positions. Every one of these positions gives rise to a set of mutually independent factors, but at most only one of the positions can claim final objective validity for the factors which it represents, the remaining positions yielding factors which depend on the particular battery used. Nevertheless, for

any one of this infinite number of positions it is possible, theoretically at least, to get together a suitable collection of tests, measuring the factors represented by it and displaying simple structure. By way of illustration, let us assume a battery of tests of ability involving only two independent objective common factors, say, for the purpose of argument, a geometric and a verbal factor. We analyze this battery by the centroid method, obtaining a general positive factor and a bipolar factor. *Ex hypothesi* these factors are not objective, but they are independent. Without rotating them we may devise a new set of tests to measure them as they stand. Two kinds of test will be necessary, one in which credit is given for all ability, whether verbal or geometric, and another in which the score depends wholly on the ratio of verbal to geometric ability. If simple structure is inserted into this new battery, suitable orthogonal rotation of the resulting centroid matrix will discover it again. But *ex hypothesi* the factors thus discovered, although independent, are not objective.

Simple structure, however, is not necessarily orthogonal, and, despairing of simple structure by orthogonal means, investigators are sometimes content to take it in an oblique form. In this way, however, no new relevant considerations are introduced, and it is hardly necessary to point out in detail by an argument similar to that which we have just given, that objectivity may still be lacking.

To sum up, the most that could be claimed for the discovery of simple structure, if the data were free from error, would be that the factors which have been taken out are those which have been put in. No claim to objectivity can be made unless there is reason to believe that the factors as put in were themselves objective. Before proceeding, however, to consider how objectivity is to be reached, a further point requires attention, viz., the presence of error in the data.

Error and Statistical Confirmation

Simple structure, if it is actually present, will normally be obscured to some extent by error; consequently, if we have reason to believe that the data will reveal simple structure, and if we accept as a hypothesis that certain factors are present, then we are entitled to claim that the results are in harmony with our hypothesis and to a certain extent confirm it, if they do not differ from the theoretic expectation by more than can be reasonably accounted for by random error. But what we are concerned with at the moment is not a hypothesis of this kind. It is rather the assumption, which appears to be made in certain quarters, that we are entitled to look for simple structure in complete ignorance of the factors which enter into it.

This attitude suffers from the disadvantage that no reason can be given for the assumption that the factors form a simple structure, for *ex hypothesi* the factors are entirely unknown. But on the other hand it has the advantage that *any* simple structure will meet its simple needs. Indeed an infinite choice of solutions lies before it, and it will claim confirmation if the final results do not differ from any one of an infinite number of arrangements by more than the theoretic probable error. What the actual probability of such a result may be no one can say, but an analogy may be helpful. On the other hand, we may compare the concrete hypothesis that certain given factors are arranged in a simple structure with the assumption that a correlation coefficient has a particular value, such as zero. In both cases, if a probable error can be calculated, we can discover the probability that an observed deviation from the theoretic value is due to error. On the other hand, we may compare the assumption of simple structure as such, i.e., of any simple structure irrespective of the factors, to the attitude of mind which is prepared to accept not one definite value for the correlation coefficient, but any one of a number scattered over the whole range, e.g., any value with zero in the second decimal place. "Confirmation" now, if we use the old probable error, is of course very much easier to obtain. The requirement is no longer that we should be within a given distance of a single point, but that we should be within that distance of any one, no matter which, of a number of different points. The probability that "confirmation" will be obtained is of course very much larger in the second than in the first case, but the probability that the underlying hypothesis is correct has not really increased correspondingly. Similarly, when we declare our intention to be satisfied with any simple structure, involving any factors, we have increased very greatly, although to an unknown extent, the number of ways in which our hypothesis can apparently be confirmed.

The nature of confirmation in a statistical argument is not always understood. A result can be regarded as confirmation only when it is consistent with one hypothesis and inconsistent with others. In the case of a numerical result the value of confirmation depends on the ratio of the range within which the result must fall if a hypothesis is correct, to the total available range. If this ratio is small confirmation is valuable; if it is large it may be worthless. The attitude which accepts any simple structure increases the magnitude of this ratio, and consequently also reduces the value of any result to which it leads.

The probability that a result concordant with a specific concrete simple structure should be obtained by chance under ordinary circumstances is small, and in general may be neglected. But the prob-

ability that a chance arrangement will approximate to any form of simple structure is not at all negligible, particularly if the probable errors are reasonably high, the factor loadings low, and the factors numerous. It may thus be taken for granted that anyone looking for a simple structure in a centroid matrix will find it far more frequently than it exists in reality.

The Principle of Parsimony

We are now in a position to return to a point mentioned earlier but passed over for the time being, viz., the claim that the search for simple structure is in harmony with the principle of parsimony. Here again there is danger of misunderstanding. The principle applies to the field of science as a whole, but if we employ it mechanically in limited isolated sections of the field we may fail to grasp its true significance. It is not uncommon to find that the intercorrelations of a small group of tests may be explained by one or two common factors, but that when these same tests are embodied in a larger battery, a larger number of factors becomes necessary, and, moreover, we may be forced to conclude that all of these factors are present as common factors in the tests of the original group. Thus we have shown elsewhere* that a particular battery of ten intelligence tests implied five common factors. We found that by special selection from this battery a sub-battery of five tests could be obtained, such that the intercorrelations of its variables could statistically be accounted for by one common factor. This one common factor, however, was incompatible with the simplest arrangement of factors which would account for the ten tests. If we assume the single factor, we explain the sub-battery in the simplest way, but we make the explanation of the larger battery, including the sub-battery, more complex, if not impossible.

The assumption that simple structure, or the nearest possible approach to it, should be our guide in the process of rotation will bias us towards a set of *ad hoc* common factors for each battery. Principles are thus multiplied, each test being explained by a different set of principles according to the company it keeps. Being penny wise we are pound foolish. No doubt it is sound practice, when dealing with limited groups of tests, to try the simpler hypotheses first and to restrict the number of factors as far as possible, but though we begin in this way we may not always end in it. We must remember that parsimony as a whole may require us to attribute complexity to a relatively small battery of tests, and it may be necessary to accept

* Reyburn, H. A. and Taylor, J. G. Some factors of intelligence. *Brit. J. Psychol.* (General Section), 1941, 31, 259 f.

a larger number of factors than the minimum which a centroid analysis would ordinarily yield.

The Function of Hypotheses

This brings us back to the conception of objectivity. Common factors, to be of value to us, must not be relative to the batteries of tests into which they enter and by which they are measured. And it follows from this that the criterion of the adequacy or objectivity of a factor cannot lie within the four corners of the battery itself. All that the analysis of a battery can yield is a hypothesis or the confirmation of a hypothesis. The factors provisionally accepted in the analysis of one battery must be tested by the analysis of others, and the ultimate criterion is the coherence of all the analyses into which these factors enter.

Factorial analysis thus, like other branches of knowledge which rest on an inductive basis, implies the use of hypotheses. There is of course no specific recipe for the manufacture of hypotheses. One thing, however, may be said with reasonable security: the less we know about a complex subject matter the less useful our hypotheses are likely to be; and ignorance of this kind, even if artificially induced by a neglect of what we already know, is not compensated for by mechanical methods of factor manipulation.

In constructing a hypothesis we should use all the knowledge that we already have. In a new field much of this may be of little value and hypotheses springing from it may require great modification or even rejection. But when the field has already been surveyed, particularly if quantitative methods have been used, it seems reasonable to accept as part of a hypothesis factors already identified and measured. Of course the acceptance is provisional, and the factors with which we begin may be altered or even rejected in the light of fuller knowledge. No hard and fast lines of procedure can be laid down, but one point, perhaps a minor one, may be mentioned. If two or more batteries contain a number of tests in common, a hypothesis which gives these tests concordant structural analyses in the different batteries is to be preferred, particularly if, within the limits of random error, it gives them the same factor loadings.

A New Technique

1. It is obvious that if we start from this point of view we must employ a somewhat different rotational method from that which has been developed by Thurstone. Part of his technique is to find a number of tests which do not contain a given factor, even although the

nature of that factor is not known, and by means of them to define a plane or hyperplane. In contrast with this fundamentally negative procedure, we suggest that the primary task is to discover tests which do contain a suspected or provisionally accepted factor with the highest degree of saturation, and by means of them to locate a single axis.* This is the first step, and as a check on it we may demand that the factor loadings measured on this axis should be reasonable and intelligible for all the variables in the battery and not only for those by means of which the axis is determined.

2. The next step is to rotate the whole centroid matrix in such a way that all the other axes are at right angles to one another and to the provisionally accepted one.

3. The third step is to set this factor aside. This is achieved by dropping the first column of the rotated matrix.

4. The whole process is now repeated on the reduced matrix. A second hypothetical factor is provisionally identified in the variables and an axis drawn to represent it; the same test of intelligibility is applied and the same method of rotation adopted.

5. The matrix is again reduced by the elimination of another column, and the procedure outlined above is repeated so long as any further hypothesis suggests itself.

6. In the present state of knowlege it will not infrequently happen that when a number of factors, say half a dozen, have to be identified, it is relatively easy to suggest a meaning for two or three or even four, but it is more difficult on the basis of previous knowledge to give an interpretation for the remainder. In such a case it may be useful to consider the remaining factors together, generally by graphic means, and to accept the most plausible arrangement of them.

7. When a provisional interpretation has thus been given to the factors, they should all be reconsidered together to see whether any adjustment by means of a further rotation will improve their intelligibility as a whole and their coherence with other analyses. This is the ultimate test we can apply: the more intelligible an analysis is as a whole and the more it is in conformity with the rest of our knowledge, the stronger is the claim which each of the factors in it has upon us.

8. Of course when analyses are not in harmony the fault may not lie in the new interpretation but in the old. We begin with a hypothesis based on previous knowledge; we may have to revise it.

* A mathematical account of the procedure described here is given in the appendix to this article.

The conceptions with which we begin and which may be current in the literature, may prove inadequate and may require to be recast. Normally we fit new ideas into an old framework and regard them as an extension of it, but when this proves not to be possible, the old framework has to be altered to make room for them. The point, however, need not be elaborated, as it is a commonplace of inductive science.

In conclusion one final comment is required. The failure to find satisfactory factor loadings for a given set of factors in a particular battery may be due, not to the lack of objectivity in the factors, but to the inadequacy of the data. Random error may tell against the recognition of the true position and may obscure a sound analysis. Hence a single failure to verify a hypothesis, while discouraging, is not fatal to it. The idea that a single negative instance is sufficient to discredit a theory is a doctrine of logicians rather than of working scientists who know the nature and limits of measurement.

APPENDIX

The method of rotation suggested in this article begins with the formulation of a hypothesis regarding one or other of the common factors entering into the battery, together with the assumption that certain of the variables are so heavily loaded with this factor that its axis will pass through or near their center of gravity. Let the selected variables be designated a, b, c, $\cdots$ and let $\mathbf{v}_a$, $\mathbf{v}_b$, $\mathbf{v}_c$, $\cdots$ be the corresponding rows of the centroid matrix $\mathbf{C}$. Then

$$\mathbf{v}_{\alpha} = \mathbf{v}_a + \mathbf{v}_b + \mathbf{v}_c + \cdots \tag{1}$$

defines a point on an axis passing through the centroid of a, b, c, $\cdots$, and the direction cosines of that axis are given by

$$\frac{1}{h_{\alpha}} \mathbf{v}'_{\alpha} = \mathbf{t}_1, \tag{2}$$

where $\frac{1}{h_\alpha}$ is a scalar, h^2_α being the sum of the squares of the elements in $\mathbf{v}_\alpha$. The first factor loadings are given by

$$\mathbf{C}\mathbf{t}_1 = \mathbf{f}_1 . \tag{3}$$

These factor loadings may show the hypothesis to be mistaken, and in that event it must be modified by selection of a different group of variables. But if it should be confirmed we proceed at once to the next step.

Since the solution is to be orthogonal, it is necessary to rotate the

remaining axes until they are all at right angles with the first axis and with one another. We have presented elsewhere* a method of writing an orthogonal transformation containing in one of its columns a given set of direction cosines. Let $\mathbf{T}_1$ be an orthogonal transformation in which $\mathbf{t}_1$ is the first column. Then

$$\mathbf{CT}_1 = \mathbf{F}_1 \tag{4}$$

gives an orthogonal factorial matrix of which the first column contains the first factor loadings and the remaining columns have no scientific significance. It will be shown below that it is not necessary to compute more than the first columns of $\mathbf{F}_1$, but for purposes of exposition it is assumed for the moment that it has been written in full.

The next step is to formulate a hypothesis concerning a second common factor and to select a second group of variables, $j, k, 1, \cdots$ which are assumed to be heavily loaded with this factor. Let the rows of $\mathbf{F}_1$ corresponding to these variables be designated $\mathbf{u}_j, \mathbf{u}_k, \mathbf{u}_l, \cdots$ and let

$$\mathbf{u}_\beta = \mathbf{u}_j + \mathbf{u}_k + \mathbf{u}_l + \cdots. \tag{5}$$

Now the second axis is to be at right angles to the first, and accordingly the first of its direction cosines, referred to the co-ordinate frame of $\mathbf{F}_1$, must be zero. Hence, if its position is to be determined by the variables $j, k, 1, \cdots$, with this restriction, we must write

$$\mathbf{w}_\beta = \mathbf{w}_j + \mathbf{w}_k + \mathbf{w}_l + \cdots, \tag{6}$$

which is derived by putting the first element of each of the row matrices in (5) equal to zero. The direction cosines of the second axis, referred to the frame of $\mathbf{F}_1$, will then be

$$\frac{1}{h^2_\beta}\mathbf{w}'_\beta = \mathbf{t}_2, \tag{7}$$

where h^2_β is the sum of the squares of the elements in $\mathbf{w}_\beta$, and

$$\mathbf{F}_1\mathbf{t}_2 = \mathbf{f}_2 \tag{8}$$

gives the second factor loadings.

We now present a method by which the same result may be reached with considerably less arithmetic. From (4) and (8) we have

$$\mathbf{CT}_1\mathbf{t}_2 = \mathbf{f}_2, \tag{9}$$

and it follows that if $\mathbf{t}_2$ can be derived by a more direct process it will

* Reyburn, H. A. and Taylor, J. G. Some factors of personality. *Brit. J. Psychol.* (General Section), 1939, 30, 159 ff.

be unnecessary to compute the whole of $\mathbf{F}_1$. Let $\mathbf{v}_j, \mathbf{v}_k, \mathbf{v}_l, \cdots$ be the rows of $\mathbf{C}$ corresponding to variables $j, k, l, \cdots$. Then from (4)

$$\mathbf{v}_j \mathbf{T}_1 = \mathbf{u}_j; \quad \mathbf{v}_k \mathbf{T}_1 = \mathbf{u}_k; \cdots$$

and

$$\mathbf{v}_\beta \mathbf{T}_1 = \mathbf{v}_j \mathbf{T}_1 + \mathbf{v}_k \mathbf{T}_1 + \mathbf{v}_l \mathbf{T}_1 + \cdots = \mathbf{u}_\beta . \tag{10}$$

From $\mathbf{u}_\beta$ we derive $\mathbf{w}_\beta$ by putting its first element equal to zero, from (7) we obtain $\mathbf{t}_2$, and from (9) we derive $\mathbf{f}_2$ without the necessity of first calculating $\mathbf{F}_1$.

Before locating the third axis we must write an orthogonal transformation $\mathbf{T}_2$ in which the first column has unity in the diagonal cell and the second column is $\mathbf{t}_2$, and compute the product

$$\mathbf{T}_1 \mathbf{T}_2 = \mathbf{T}_3 .$$

Then, proceeding as before, we obtain

$$\mathbf{v}_\gamma \mathbf{T}_3 = \mathbf{u}_\gamma ,$$

where $\mathbf{v}_\gamma$ is derived from the rows of $\mathbf{C}$ corresponding to the variables assumed to be mainly loaded with the third factor. Again, $\mathbf{w}_\gamma$ is derived from $\mathbf{u}_\gamma$ by putting its first two elements equal to zero. The direction cosines of the third axis, referred to the co-ordinate frame of $\mathbf{C}$, are given by

$$\mathbf{T}_3 \frac{1}{h_\gamma} \mathbf{w}'_\gamma = \mathbf{t}_3$$

and the third factor loadings are obtained from

$$\mathbf{C}\mathbf{t}_3 = \mathbf{f}_3 .$$

This process is continued until no further fruitful hypothesis suggests itself, or until it becomes more convenient to complete the rotation by graphical means.

To ensure accurate working, the following checks should be applied.

(a) An orthogonal transformation is checked if $\mathbf{T}\mathbf{T}' = \mathbf{I}$. But the method of constructing the transformations used here is such that it is not necessary to calculate every element in the product. If each diagonal element is unity, and the elements immediately adjacent to the diagonal are all zero, the remaining elements will all be zero. It should be added that in constructing orthogonal transformations when only one column is known, we have found it advisable to work to not fewer than five decimal places.

(b) The calculation of $\mathbf{u}_\beta$ in (10) is checked if

$$\mathbf{v}_\beta \mathbf{v}'_\beta = \mathbf{u}_\beta \mathbf{u}'_\beta ,$$

that is, if the sums of the squares of $\mathbf{v}_\beta$ and $\mathbf{u}_\beta$ are identical.

(c) The factorial matrix **F** is checked if

$$\mathbf{FF}' = \mathbf{CC}' = \mathbf{R} .$$

Here again it is not necessary to compute more than the diagonal elements (communalities) and the elements immediately adjacent to the diagonal. That is, we check h^2 for every variable, and also r_{12}, r_{23}, r_{34}, etc. Even if the values of h^2 are all correct, errors may be made in the recording of signs, and these are most readily detected by reconstructing from **F** the correlations between each variable and its nearest neighbours.